AD_____

Award Number: DAMD17-02-1-0135

TITLE: High Throughput Synthesis and Screening for Agents Inhibiting Androgen Receptor mediated Gene Transcription

PRINCIPAL INVESTIGATOR: Dale L. Boger, Ph.D.

CONTRACTING ORGANIZATION: The Scripps Research Institute
La Jolla, California 92307

REPORT DATE: February 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	February 2004	Annual (1 Feb 2003 - 31 Jan 2004)

4. TITLE AND SUBTITLE
High Throughput Synthesis and Screening for Agents Inhibiting Androgen Receptor mediated Gene Transcription

5. FUNDING NUMBERS
DAMD17-02-1-0135

6. AUTHOR(S)
Dale L. Boger, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
The Scripps Research Institute
La Jolla, California 92307

E-Mail: boger@scrupps.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)
Details of the year 2 progress on our grant to define a new treatment of therapeutic resistant prostate cancer is provided. This entails the high throughput synthesis of DNA binding agents related to distamycin, their screening for binding to androgen response elements using a new high throughput DNA binding screen we introduced, and the evaluation for inhibiting androgen receptor mediated gene transcription and the cell proliferation of a prostate cancer cell line.

14. SUBJECT TERMS
Prostate cancer, DNA minor groove binding agents, androgen receptor, androgen response elements

15. NUMBER OF PAGES
66

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

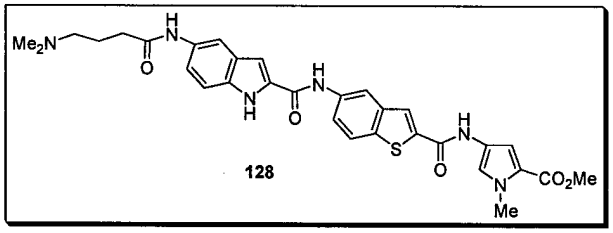
REPORT DOCUMENTATION PAGE (SF298)
(Continuation Sheet)

Revised Report.

Introduction. The objective of our work is the high throughput synthesis of DNA minor groove binding agents based on the distamycin structure and their screening for inhibition of androgen receptor mediated gene transcription, which is unregulated in chemotherapeutic resistant prostate cancer and responsible for disease progression. It represents a novel and unique new target for the treatment of relapse prostate cancer where prognosis is presently very poor.

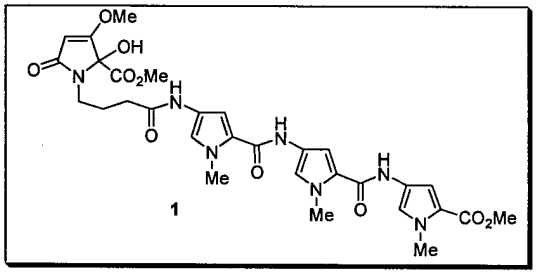
Body. The results to date in our studies are presented below and refer to Task 1–Task 4 of our proposals' statement of work.

Task 1: To synthesize combinatorial libraries and compounds. In year 2, we found that **128** potently (IC_{50} = 8 nM) and selectively (*ca.* 100-fold) inhibits androgen receptor mediated gene transcription in the assay we detailed in the proposal, but only if it is delivered artificially into the cell using liposomes (publications 3 and 5). Thus, the strategy outlined can and does identify candidate compounds useful for potential treatment of therapeutic resistant prostate cancer. However, we have not yet identified derivatives that enter cells better than **128** or inhibit androgen receptor mediated gene transcription (in a cellular functional assay) in the absence of liposomal delivery.



128

In year 2, we also examined the feasibility of removing the charged group (which inhibits cell penetration) and attaching a novel reversible alkylating agent capable of increasing (replacing) the lost DNA binding affinity that results from removing the charged group (publication 4). In these studies, we did not identify such derivatives among the prototypical compounds prepared (e.g. **1**). That is, the attachment of such a reversible alkylating group did not increase DNA binding affinity (i.e., no apparent alkylation). This work complements our continuing efforts to modulate cellular uptake of such derivatives by altering the pK_a of the terminal amine (Year 1 and continuing studies).



1

In year 2, we also continued our preparation of additional DNA binding libraries to discover new leads (see Figure 2b of year 2 report). The preparation of this second generation library of 4,394 members was completed (complementing our initial library of 2,640 compounds).

For this purpose, we also surveyed a series of building blocks that could be part of this library and examined their ability to bind DNA in our assay entailing the displacement of ethidium bromide from hairpin DNA's. The results of this work were published in year 1 and the paper detailing this work was previously provided in the year 1 appendix[1] (C.R. Woods, N. Faucher, B. Eschgfaller, K.W. Bair, and D. L. Boger,

Synthesis and DNA binding properties of saturated distamycin analogues, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2647–2650).

Consequently, our progress on Task 1 is now completed.

Task 2: To identify libraries that have antitumor activity for prostate cancer cell lines and have high affinity binding to AREs.

- The cytotoxic assay for the prostate cancer cell line, LNCaP, was set up and validated (year 1).
- The first library of 2,640 compounds was tested in this cell line versus the control L1210 cell line (this year 2).
- The screening of our second library of 4,394 members in this cell line has been initiated (Year 2 and 3).
- The screening of the first library of 2,640 compounds against the ARE hairpin DNAs has been performed (year 1).

Once completed for both libraries, the data for the two assays will be correlated to identify active constituents that display activity against the prostate cancer cell line and bind the ARE consensus sequence. Already, six new lead structures have emerged from the library 1 screening (year 2).

Task 3: Define DNA binding selectivity of identified new leads.

This work is well underway and has been completed (screening and analysis) for the six new leads that emerged from the 1st generation library of 2,640 compounds. We will be completing a similar study of the leads that emerge from the 4,394 compounds in the next 2–4 months using the assay we developed (Boger, et al. *J. Am. Chem. Soc.* **2001**, *123*, 5878–5891).[2]

Task 4: Determine the inhibitory effect on AR mediated gene transcription and transactivation.

As indicated in the original proposal, work on this task will be conducted at the final stages of the grant (year 3). The reported constructs have been prepared and are ready to implement.

Key Research Accomplishments.

- Development and implementation of a novel solution-phase approach to the preparation of libraries of sequence selective DNA binding compounds (high throughput synthesis).[3]
- Development of a novel high throughput screen for establishing DNA binding selectivity or affinity. This includes the introduction of the first high throughput screen for a defined DNA sequence (i.e., androgen response element) that can control aberrant gene transcription.[2] We recently reviewed our efforts on the development of this screen and highlighted its application to inhibition of androgen receptor mediated gene transcription (publication 5).[2c]
- Definition of previously unexamined features of distamycin responsible for its DNA binding affinity (publication 1).[1]
- Highlighted the feasibility of discovering inhibitors of androgen mediated gene transcription with sequence selective minor groove DNA binding compounds, albeit presently requiring liposomal delivery of the compounds into a cell (publications 2 and 3).[4]

Reportable Outcomes.

Publications

1. C. R. Woods, N. Faucher, B. Eschgfaller, K. W. Bair, and D. L. Boger, Synthesis and DNA binding properties of saturated distamycin analogues, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2647–2650.

This grant period:

2. D. L. Boger, Solution-phase synthesis of combinatorial libraries designed to modulate protein–protein or protein–DNA interactions, *Bioorg. Med. Chem.* **2003**, *11*, 1607–1613.
3. D. L. Boger, J. Desharnais, and K. Capps, Solution-phase combinatorial libraries: modulating cellular signaling by targeting protein–protein or protein–DNA interactions, *Angew. Chem. Int. Ed.* **2003**, *42*, 4138–4176.
4. B. K. S. Yeung and D. L. Boger, Synthesis of isochrysohermidin–distamycin hybrids, *J. Org. Chem.* **2003**, *68*, 5249–5253.
5. W. C. Tse and D. L. Boger, A fluorescent intercalator displacement (FID) assay for establishing DNA binding selectivity and affinity, *Acc. Chem. Res.* **2004**, *37*, 61–69.

Conclusions. The work is progressing well and requires no changes in future work. The importance of the work includes not only the potential development of a treatment for relapse (resistant) prostate cancer, but it defines a new approach to treating diseases arising from abberant gene transcription and, importantly, provides the first scientific tools to approach this problem (high throughput synthesis and screening technology for DNA binding compounds).

References.

1. C. R. Woods, N. Faucher, B. Eschgfaller, K. W. Bair, and D. L. Boger, Synthesis and DNA binding properties of saturated distamycin analogues, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2647–2650.
2. (a) D. L. Boger, B. E. Fink, S. R. Brunette, W. C. Tse, and M. P. Hedrick, A simple, high resolution method for establishing DNA binding affinity and sequence selectivity, *J. Am. Chem. Soc.* **2001**, *123*, 5878–5891. (b) D. L. Boger and W. C. Tse, Thiazole orange as the fluorescent intercalator in a high resolution FID assay for determining DNA binding affinity and sequence selectivity of small molecules, *Bioorg. Med. Chem.* **2001**, *9*, 2511–2518. (c) W. C. Tse and D. L. Boger, A fluorescent intercalator displacement (FID) assay for establishing DNA binding selectivity and affinity, *Acc. Chem. Res.* **2004**, *37*, 61–69.
3. D. L. Boger, B. E. Fink, and M. P. Hedrick, Total synthesis of distamycin A and 2640 analogues: a solution-phase combinatorial approach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity, *J. Am. Chem. Soc.* **2000**, *122*, 6382–6394.
4. (a) D. L. Boger, Solution-phase synthesis of combinatorial libraries designed to modulate protein–protein or protein–DNA interactions, *Bioorg. Med. Chem.* **2003**, *11*, 1607–1613. (b) D. L. Boger, J. Desharnais, and K. Capps, Solution-phase combinatorial libraries: modulating cellular signaling by targeting protein–protein or protein–DNA interactions, *Angew. Chem. Int. Ed.* **2003**, *42*, 4138–4176.

Appendix. Attached

APPENDIX

Pergamon

SCIENCE DIRECT®

Bioorganic & Medicinal Chemistry 11 (2003) 1607–1613

BIOORGANIC & MEDICINAL CHEMISTRY

Perspective

Solution-Phase Synthesis of Combinatorial Libraries Designed to Modulate Protein–Protein or Protein–DNA Interactions

Dale L. Boger*

Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, CA 92037, USA

Received 6 December 2002; accepted 6 December 2002

Abstract—A short personal perspective on the development of an approach to the solution-phase synthesis of combinatorial libraries for modulating cellular signaling by inhibiting, promoting, or mimicking protein–protein or protein–DNA interactions is provided.

Over the course of the last several years, we have developed technology for the solution-phase preparation of combinatorial libraries with several key applications in mind. Because this work was necessarily published at different intervals across a range of chemical or biological journals, the strategic decisions we made in pursuing this work and the intimate relationship between the technology development and the intended applications is not likely to be clear to the casual reader. Consequently, I would like to take the opportunity of this Perspective Articles series to summarize our contributions to solution-phase library synthesis highlighting its development in the context of our intended applications for modulating cellular signaling through inhibiting, promoting, or mimicking protein–protein or protein–DNA interactions. Despite the prevailing bias that such targets might not prove viable for small molecule intervention, the screening of our libraries against each such target has provided the first small molecule modulators of the protein–protein or protein–DNA interaction and validated the targets for small molecule therapeutic intervention.

Solution-Phase Combinatorial Library Synthesis

Combinatorial chemistry has undergone rapid development and has provided a new paradigm for drug discovery. As a consequence of its extension from peptide and oligonucleotide synthesis, most approaches have relied on solid-phase synthesis techniques. A complement to adapting solution-phase chemistry to polymer-supported combinatorial synthesis is the development of protocols for solution-phase combinatorial synthesis. We introduced a simple protocol that permits the multistep synthesis of chemical libraries employing liquid–liquid and liquid–solid extractions to remove unreacted starting materials, reagents and reagent by-products, providing the purified product (>95% pure) irrespective of the reaction efficiency (Fig. 1).[1–34] It has been implemented in formats for the synthesis of individual compounds[1–4] (1000-member libraries), modest sized libraries composed of small mixtures (1000- to 10,000-member libraries, 10–50 compounds/mixture),[6] or combinatorially assembled to provide large libraries including positional scanning or deletion synthesis libraries[8,13,21,34] (25,000- to 1,000,000-member libraries, 100–28,000 compounds/mixture).[7,8] This allows the protocol to be adopted in a format compatible with any screening objective. Thus, its implementation is convenient for either lead discovery or lead optimization and produces the library members on a scale (5–150 mg) that allows their repeated use in screening without resynthesis. It is this latter feature along with its technically nondemanding implementation that we consider its greatest attributes. We presently have 40,000 compounds in the small (1000-member) and medium-sized (1000–10,000-member) libraries and efforts are ongoing to expand this to approximately 1,000,000 compounds. Such libraries have been prepared that interfere with (antagonists) or mimic (agonists) extracellular or intracellular protein–protein interactions,[35–37] inhibit intracellular enzymes, or modulate protein–DNA interactions. The chemistry is applicable to non-natural and natural product scaffolds, and cyclic (depsi)peptides and possesses a scope that exceeds what one might initially imagine based on its simplicity.

*Corresponding author. Tel.: +1-858-784-7522; fax: +1-858-784-7550; e-mail: boger@scripps.edu

0968-0896/03/$ - see front matter
doi:10.1016/S0968-0896(03)00031-2

Highlights of the methodology developed for the preparation of chemical libraries include:

- A general solution-phase technique for the high-throughput synthesis of libraries in a variety of formats (individual compounds, small mixtures, large mixtures, positional scanning libraries, deletion synthesis libraries) was introduced that is applicable to both lead identification and lead optimization.
- Liquid–liquid or liquid–solid (ion-exchange resin) extractions were developed.[1,2,11]
- Immobilized reagents for use in the solution-phase synthesis of libraries were introduced.[2,10–12]
- The first use of size exclusion chromatography (filtration) for the high-throughput purification and isolation of chemical libraries was disclosed.[11]
- The first application of an intermolecular olefin metathesis reaction in the synthesis of chemical libraries was disclosed.[5,7–9]
- The approach is applicable to the multistep synthesis of selected natural products and this was highlighted with a 2640-membered distamycin library,[20] 132-membered CC-1065 library,[24] 40-membered HUN-7293 library,[18] and a prototype triostin A synthesis.[22]
- Deletion synthesis deconvolution of mixture libraries was introduced and assessed.[8]
- The performance of positional scanning deconvolution was and continues to be compared against deletion synthesis deconvolution, small mixture screening, and individual compound screening.[8,13,21,34]
- Convergent versus divergent strategies of library synthesis were introduced and developed (only applicable to solution-phase, not solid-phase, synthesis).[7–12]
- Agonists from antagonists design concepts implemented.[32]
- A rapid, high-throughput, high-resolution screen for DNA binding affinity and selectivity applicable for use in assaying libraries was developed (FID assay).[20,23,25–27]
- Current archived chemical libraries contain >40,000 compounds that are available for continued screening in new assays.

The approach avoids the disadvantages of solid-supported synthesis including its more restrictive scale, the required functionalized substrates and solid supports, compatible spacer linkers, and the requirements for orthogonal attachment/detachment chemistries. It does not require specialized protocols for monitoring each step of multistep syntheses, allows the purification of intermediates, and provides the final pure products directly for use in binding or functional assays. We described extensions of these studies for generating symmetrical or unsymmetrical chemical libraries suitable for probing receptor and protein homo- and heterodimerization events (Fig. 2). Thus, the preparation and dimerization linkage of iminodiacetic acid diamides can be conducted in a reaction sequence that requires only three steps. In addition to the multiplication of the diversity that arises through the combinatorial dimerization linkage of the iminodiacetic acid diamides, the solution-phase synthesis of the intermediates permits their direct linkage which would be precluded by solid-phase synthesis techniques. As such, the strategy is uniquely suited for taking advantage of such dimerization (convergent) strategies utilizing a limited number of synthetic steps. This modular approach to the generation of libraries is especially well-suited for the discovery of antagonists or agonists of receptor and protein homo- and heterodimerization. Simple binders (i.e., **3**) can serve as antagonists of ligand-induced receptor or protein dimerization. Covalently linked symmetrical dimers (i.e., **4**) can be used to promote receptor or protein homodimerization whereas unsymmetrical dimers can be utilized to promote receptor or protein heterodimerization. Thus, both antagonists and agonists may be developed depending on the therapeutic application.[41,45]

Target Protein–Protein Interactions Mediating Cellular Signaling

Cell growth, differentiation, migration, and apoptosis are regulated in part by growth factors or cytokines. These factors are unable to penetrate the cell membrane and exert their effects by binding to cell surface receptors. In many instances, such receptors are activated by ligand-induced dimerization or oligomerization.[38–44] In addition, several components of the intracellular signal transduction pathways are also regulated by dimerization. For instance, certain cytoplasmic signal transduction molecules dimerize after activation, and the active form of a transcription factor is often a dimer.[42–45] Thus, protein dimerization has emerged as a general mechanism for the initiation and downstream regulation of signal transduction. The targets we are addressing constitute prototypical examples of these events in signal transduction. Thus, the targets being pursued were chosen not only for their therapeutic importance, but also because each constitutes a distinct stage for modulating cellular signaling by controlling protein–protein interactions.

Effective inhibitors of angiogenesis (new blood vessel growth) and tumor growth have been discovered that act by disruption of the binding of matrix metalloproteinase 2 (MMP2) to the cell surface integrin $\alpha_v\beta_3$ validating the target for therapeutic intervention for the treatment of cancer and providing the first small molecule lead structures (Fig. 3).[30,31] The first small molecule inhibitors of Myc/Max dimerization have been discovered that act at the bHLHZip dimerization interface enlisting a novel FRET assay that were shown to disrupt the binding of this oncogenic transcription factor to DNA and to inhibit Myc induced conversion of normal to transformed fibroblasts validating an important small molecule target for the treatment of cancer (Fig. 4).[33] Promising inhibitors of LEF-1/β-catenin mediated gene transcription have been identified in a luciferase reporter assay (TOPFLASH).[29] These inhibitors and their mechanism of action (inhibition of LEF-1/β-catenin

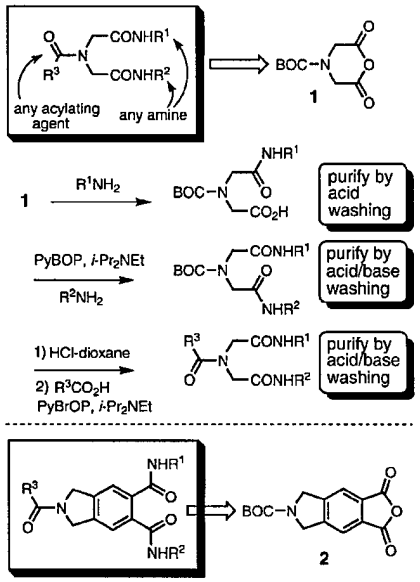


Figure 1. Solution-phase synthesis of libraries.

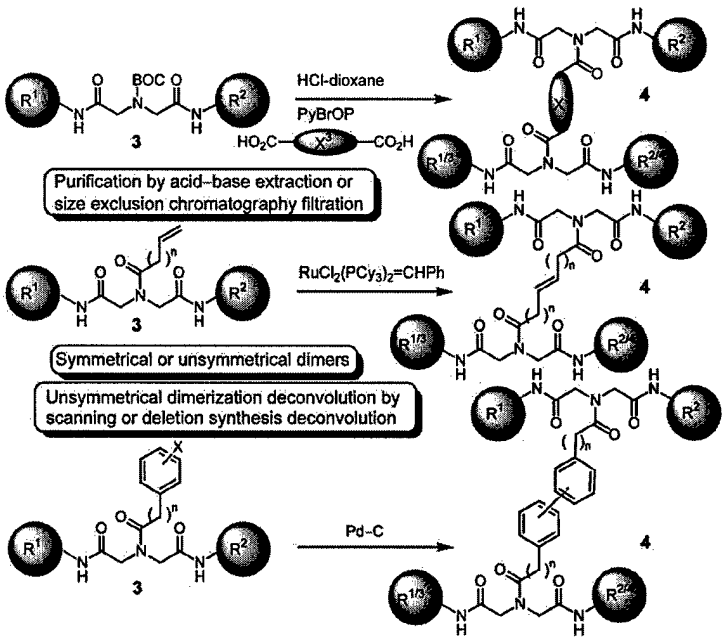


Figure 2. Homo- or heterodimerization of libraries for protein–protein dimerization.

binding) are being more fully explored with the development of a LEF-1/β-catenin FRET assay. The first inhibitors of the binding of the signaling adapter protein Paxillin to the short cytoplasmic tail of the integrin α_4 ($\alpha_4\beta_1$ = VLA4, Very Late Antigen 4) have been discovered and shown to inhibit cell migration validating a new target for the treatment of chronic inflammatory diseases, asthma, and multiple sclerosis (Fig. 5).[34] Finally, the remarkable discovery of relatively small erythropoietin (EPO) agonists that act by promoting homodimerization of the cell surface receptor (Fig. 6) has been made in an approach that could provide conventional therapeutic replacements for the recombinant human protein (7 billion/yr market) used for the treatment of anemias resulting from cancer, AIDS and other clinical disease states.[32]

The targets were carefully chosen not only for their therapeutic importance, but such that each represents a different prototypical extracellular or intracellular signaling event involving protein–protein or protein–DNA interactions:

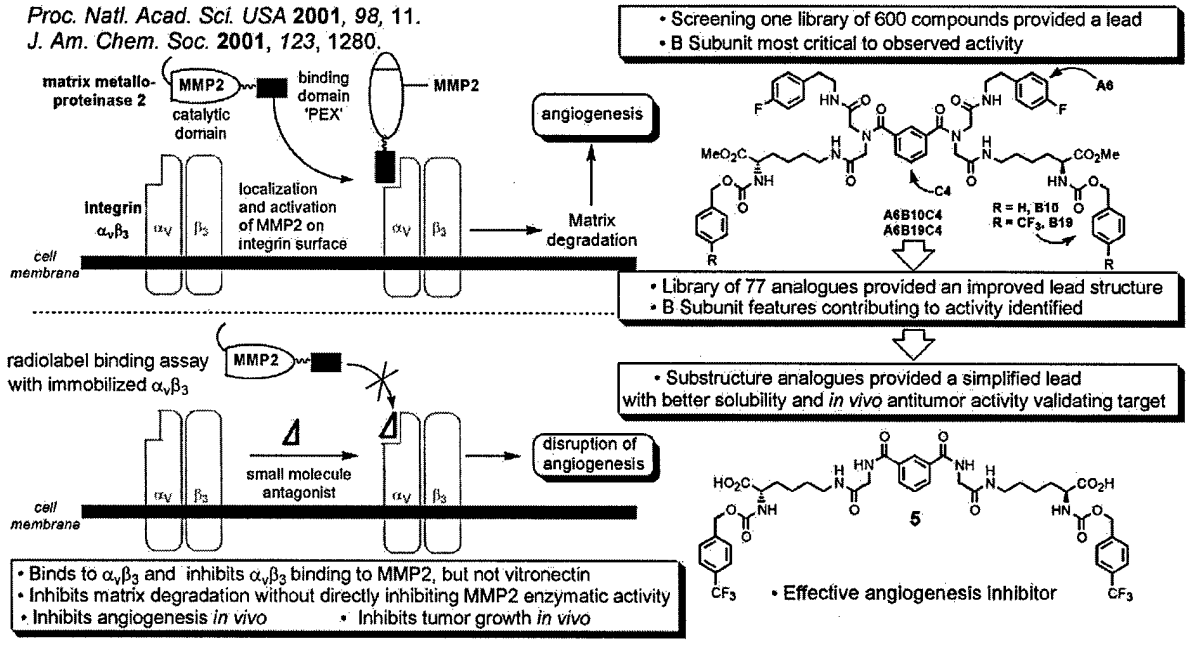


Figure 3. Disruption of angiogenesis by blocking MMP2 binding to integrin $\alpha_v\beta_3$.

- Promote a cell surface receptor homodimerization required for activation (EPO agonist)
- Inhibit an extracellular cell surface integrin–protein interaction (MMP2/$\alpha_v\beta_3$).
- Inhibit an intracellular cell surface integrin–protein interaction (Paxillin/$\alpha_4\beta_1$).
- Inhibit the intracellular heterodimerization of a transcription factor (Myc–Max, LEF-1/β-catenin).
- Inhibit a protein–DNA interaction of a transcription factor (LEF-1/β-catenin).

Protein–DNA Interactions: High-Throughput Synthesis and Screening Applied to the Discovery of Biologically Active DNA Binding Agents

We have also described the high-throughput synthesis and screening of DNA binding compounds[19–28] that are related to our interests in understanding and exploiting their properties. The approach integrates the solution-phase techniques for the synthesis of libraries with a technique we recently introduced for rapid high-throughput screening for DNA binding affinity or sequence selectivity (Fig. 7).[20,23] These techniques can be combined to rapidly explore and define the structural features responsible for the sequence selective DNA binding properties of known agents, to discover new paradigms for small molecule recognition of DNA (new bp codes), and to screen for compounds that selectively target consensus sequences of transcription factors for controlling aberrant gene transcription (i.e, LEF-1/β-catenin).

Highlights of these efforts to date include:

- Development of a high resolution, high throughput FID assay for assessing DNA binding affinity or selectivity.[20,23,25]

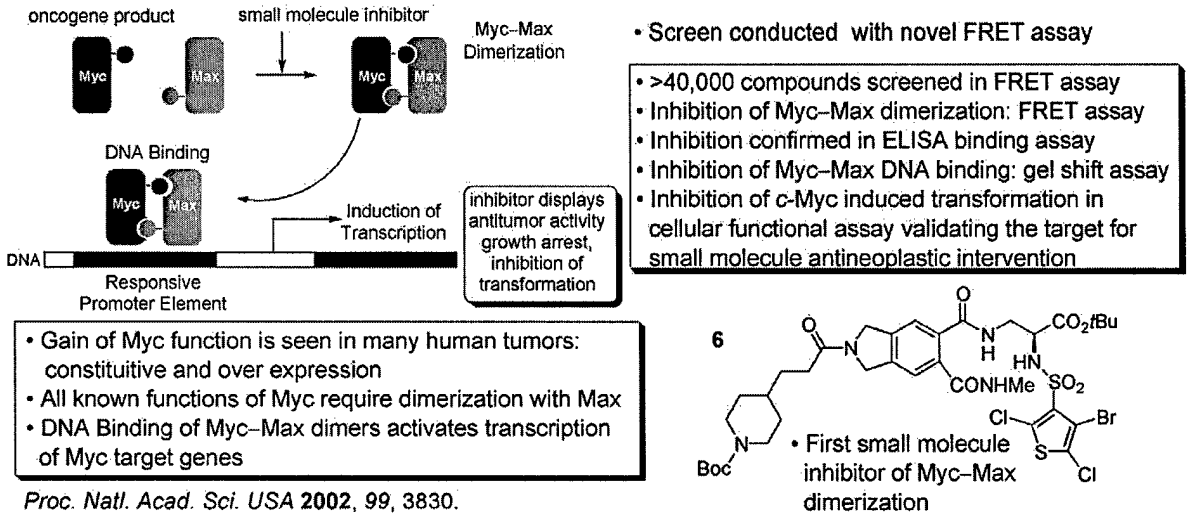


Figure 4. Inhibition of Myc–Max (transcription factor) heterodimerization and abberant gene transcription: antitumor target validation.

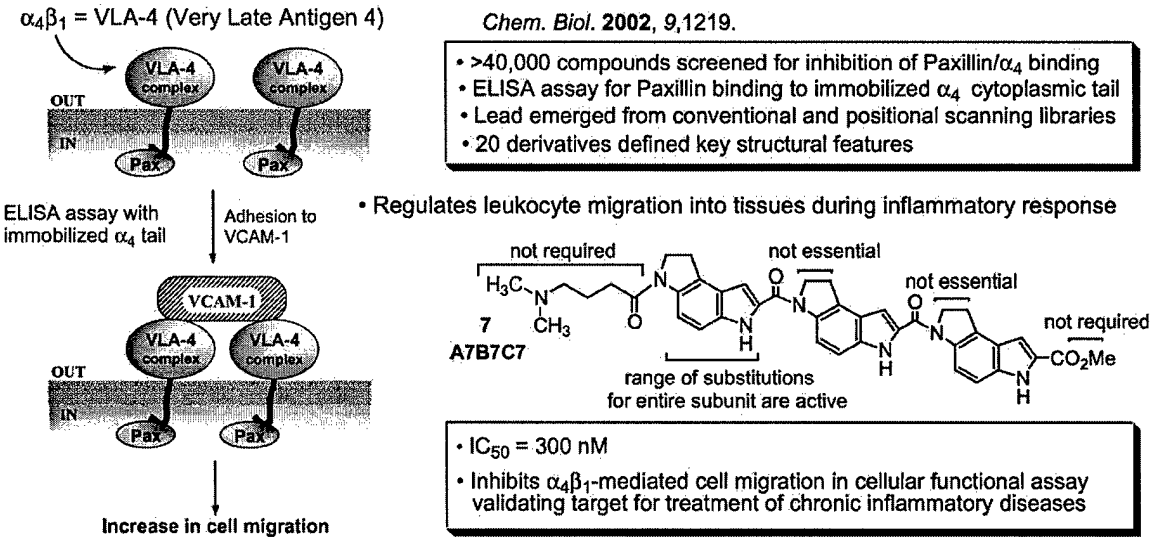


Figure 5. Inhibition of cell migration by blocking intracellular Paxillin/α_4 binding.

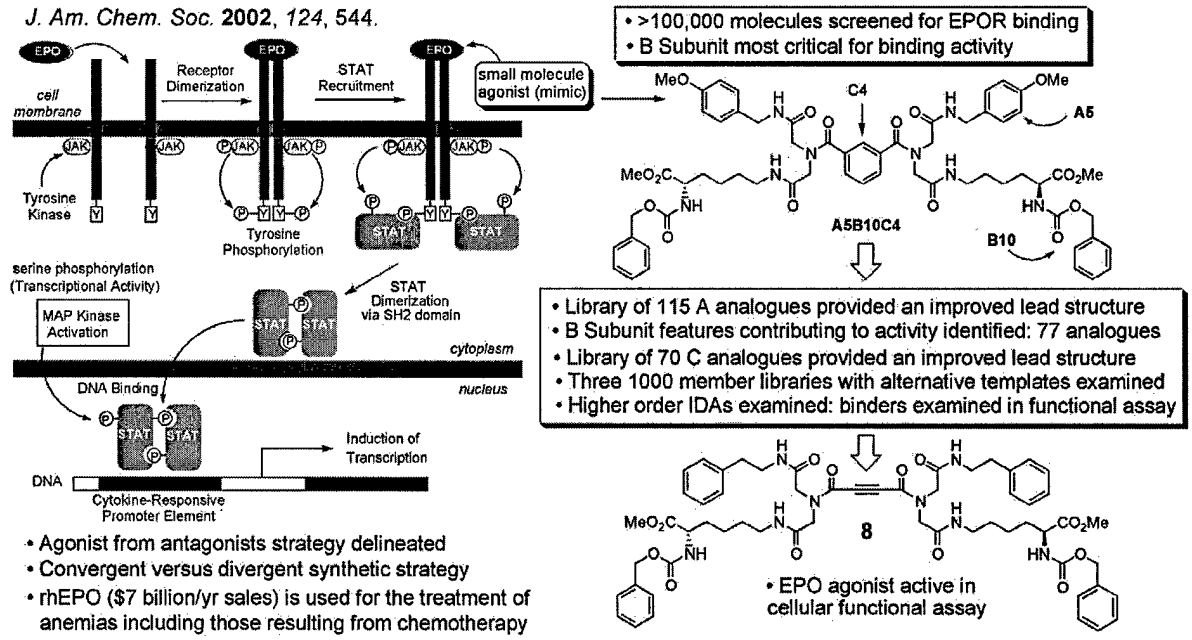


Figure 6. EPO agonists (mimics) that function by promoting EPOR dimerization.

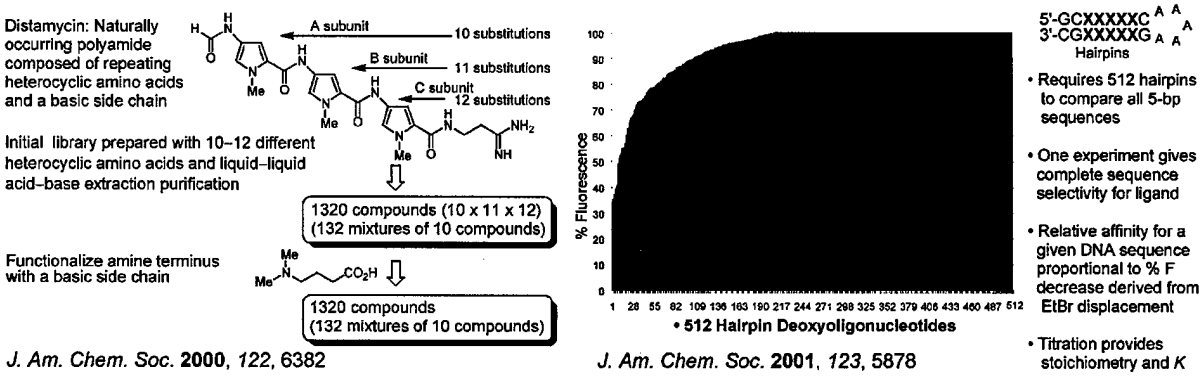


Figure 7. Solution-phase synthesis of DNA minor groove binding ligands and FID assay for DNA binding affinity and selectivity.

- Disclosure of some of the first libraries of DNA binding compounds (distamycin, CC-1065).[20–22,24]
- Discovery of a general hairpin versus extended DNA binding of a substituted β-alanine linked polyamide.[26]
- The first characterization of cooperative extended 2:1 side-by-side parallel (vs antiparallel) DNA binding with a novel class of iminodiacetic acid (IDA) linked polyamides.[27]

Acknowledgements

We gratefully acknowledge the financial support of the National Institutes of Health which has supported various aspects of our work (protein–protein interactions, CA78045; DNA binding compounds, CA41986; cyclic (depsi)peptides, CA41101; angiogenesis inhibitors, EY14174), the Department of Army (DAMD17-02-1-0135), and The Skaggs Institute for Chemical Biology.

References and Notes

1. Boger, D. L.; Tarby, C. M.; Caporale, L. H. Generalized dipeptidomimetic template: solution-phase parallel synthesis of combinatorial libraries. *J. Am. Chem. Soc.* **1996**, *118*, 2109.
2. Cheng, S.; Comer, D. D.; Williams, J. P.; Boger, D. L. Novel solution-phase strategy for the synthesis of chemical libraries containing small organic molecules. *J. Am. Chem. Soc.* **1996**, *118*, 2567.
3. Cheng, S.; Tarby, C. M.; Comer, D. D.; Williams, J. P.; Caporale, L. H.; Boger, D. L. A solution-phase strategy for the synthesis of chemical libraries containing small organic molecules: a universal and dipeptide mimetic template. *Bioorg. Med. Chem.* **1996**, *4*, 727.
4. Review: Tarby, C. M.; Cheng, S.; Boger, D. L. Solution phase strategy for the synthesis of chemical libraries containing small organic molecules: a general depeptide mimetic and a flexible universal template, In *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery*; Chaiken, I. M., Janda, K. D., Eds.; ACS: Washington, 1996, pp 81–98.
5. Boger, D. L.; Chai, W.; Ozer, R. S.; Andersson, C.-M. Solution-phase combinatorial synthesis via the olefin metathesis reaction. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 463.

6. Boger, D. L.; Ozer, R. S.; Andersson, C.-M. Generation of symmetrical compound libraries by solution-phase combinatorial chemistry. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1903.
7. Boger, D. L.; Chai, W. Solution-phase combinatorial synthesis: convergent multiplication of diversity via the olefin metathesis reaction. *Tetrahedron* **1998**, *54*, 3955.
8. Boger, D. L.; Chai, W.; Jin, Q. Multistep convergent solution-phase combinatorial synthesis and deletion synthesis deconvolution. *J. Am. Chem. Soc.* **1998**, *120*, 7220.
9. Boger, D. L.; Ducray, P.; Chai, W.; Jiang, W.; Goldberg, J. Higher order iminodiacetic acid libraries for probing protein–protein interactions. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2339.
10. Boger, D. L.; Goldberg, J.; Jiang, W.; Chai, W.; Ducray, P.; Lee, J. K.; Ozer, R. S.; Andersson, C.-M. Higher order iminodiacetic acid libraries for probing protein–protein interactions. *Bioorg. Med. Chem.* **1998**, *6*, 1347.
11. Boger, D. L.; Goldberg, J.; Andersson, C.-M. Solution phase combinatorial synthesis of biaryl libraries employing heterogeneous conditions for catalysis, isolation and size exclusion chromatography for purification. *J. Org. Chem.* **1999**, *64*, 2422.
12. Boger, D. L.; Jiang, W.; Goldberg, J. Convergent solution-phase synthesis of combinatorial libraries through rigid biaryl or diacetylene couplings. *J. Org. Chem.* **1999**, *64*, 7094.
13. Boger, D. L.; Lee, J. K.; Jin, Q. Two comparisons of the performance of positional scanning and deletion synthesis for the identification of active constituents in mixture combinatorial libraries. *J. Org. Chem.* **2000**, *65*, 1467.
14. Boger, D. L., Solution-Phase Synthesis of Combinatorial Libraries: To Bead or Not to Bead; In *Perspective Articles on the Utility and Application of Solution-Phase Combinatorial Chemistry*; Baldino, C. M., Eds.; *Comb. Chem.* **2000**, *2*, 89.
15. Review: Boger, D. L.; Goldberg, J. Multip-Step Solution-Phase Combinatorial Chemistry; In *Combinatorial Chemistry: A Practical Approach*; Fenniri, H. Ed.; Oxford University Press: Oxford, 2000; pp 303–326.
16. Boger, D. L.; Chen, Y.; Foster, C. A. Synthesis and evaluation of aza HUN-7293. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1741.
17. Review: Boger, D. L.; Goldberg, J. Cytokine receptor dimerization and activation: prospects for small molecule agonists. *Bioorg. Med. Chem.* **2001**, *9*, 557.
18. Chen, Y.; Bilban, M.; Foster, C. A.; Boger, D. L. Solution-phase parallel synthesis of a pharmacophore library of HUN-7293 analogues: a general chemical mutagenesis approach to defining structure–function properties of cyclic (depsi)peptides. *J. Am. Chem. Soc.* **2002**, *124*, 5431.
19. Boger, D. L.; Fink, B. E.; Hedrick, M. P. A new class of highly cytotoxic diketopiperazines. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1019.
20. Boger, D. L.; Fink, B. E.; Hedrick, M. P. Total synthesis of distamycin A and 2640 analogs: a solution-phase combinatorial apporach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity. *J. Am. Chem. Soc.* **2000**, *122*, 6382.
21. Boger, D. L.; Dechantsreiter, M. A.; Ishii, T.; Fink, B. E.; Hedrick, M. P. Assessment of solution phase positional scanning libraries based on distamycin A for the discovery of new DNA binding agents. *Bioorg. Med. Chem.* **2000**, *8*, 2049.
22. Boger, D. L.; Lee, J. K. Development of a solution-phase synthesis of minor groove binding bis-intercalators based on triostin A suitable for combinatorial synthesis. *J. Org. Chem.* **2000**, *65*, 5996.
23. Boger, D. L.; Fink, B. E.; Brunette, S. R.; Tse, W. C.; Hedrick, M. P. A simple, high resolution method for establishing DNA binding affinity and sequenxe selectivity. *J. Am. Chem. Soc.* **2001**, *123*, 5878.
24. Boger, D. L.; Schmitt, H.; Fink, B. E.; Hedrick, M. P. Parallel synthesis and evaluation of 132 (+)-1,2,3,9a-tetrahydrocyclopropa[*c*]benz[*e*]indol-4-one (CBI) analogues of CC-1065 and the duocarmycins defining the contribution of the DNA biding domain. *J. Org. Chem.* **2001**, *66*, 6654.
25. Boger, D. L.; Tse, W. C. Thiazole orange as the fluorescent intercalator in a high resolution FID assay for determining DNA binding affinity and sequence selectivity of small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2511.
26. Woods, C. R.; Ishii, T.; Wu, B.; Bair, K. W.; Boger, D. L. Hairpin versus extended DNA binding of a substituted β-alanine linked polyamide. *J. Am. Chem. Soc.* **2002**, *124*, 2148.
27. Woods, C. R.; Ishii, T.; Boger, D. L. Synthesis and DNA binding properties of iminodiactic acid (IDA) linked polyamides: characterization of cooperative extended 2:1 side-by-side parallel binding. *J. Am. Chem. Soc.* **2002**, *124*, 10676.
28. Woods, C. R.; Faucher, N.; Eschgfaller, B.; Bair, K. W.; Boger, D. L. Synthesis and DNA binding properties of saturated distamycin analogues. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2647.
29. Boger, D. L.; Goldberg, J.; Satoh, S.; Cohen, S. B.; Vogt, P. K. Non amide-based combinatorial libraries derived from N-BOC iminodiacetic acid: solution-phase synthesis of a 150-membered piperazinone library with activity against LEF-1/β-catenin mediated transcription. *Helv. Chim. Acta* **2000**, *83*, 1825.
30. Silletti, S.; Kessler, T.; Goldberg, J.; Boger, D. L.; Cheresh, D. A. Selective disruption of MMP2 binding to integrin $\alpha_v\beta_3$ by a novel organic molecule inhibitors angiogenesis and tumor growth in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 119.
31. Boger, D. L.; Goldberg, J.; Silletti, S.; Kessler, T.; Cheresh, D. A. Identification of a novel class of small-molecule anti-angiogenic agents through screening of combinatorial libraries which function by inhibiting the binding and localization of proteinase MMP2 to integrin $\alpha_v\beta_3$. *J. Am. Chem. Soc.* **2001**, *123*, 1280.
32. Goldberg, J.; Jin, Q.; Ambroise, Y.; Satoh, S.; Desharnais, J.; Capps, K.; Boger, D. L. Erythropoietin mimetics derived from solution phase combinatorial libraries. *J. Am. Chem. Soc.* **2002**, *124*, 544.
33. Berg, T.; Cohen, S. B.; Desharnais, J.; Sonderegger, C.; Maslyar, D. J.; Goldberg, J.; Boger, D. L.; Vogt, P. K. Small molecule antagonists of Myc/Max dimerization inhibit Myc-indiced transformation of chicken embryo fibroblasts. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3830.
34. Ambroise, Y.; Yuspan, B.; Ginsberg, M. H.; Boger, D. L. Inhibitors of cell migration that inhibit intracellular paxillin/α_4 binding: a well documented use of positional scanning libraries. *Chem. Biol.* **2002**, *9*, 1219.
35. Toogood, P. L. Inhibition of protein–protein association by small molecules: approaches and progress. *J. Med. Chem.* **2002**, *45*, 1543.
36. Cochran, A. G. Antagonists of protein–protein interactions. *Chem. Biol.* **2000**, *7*, R85.
37. Zutshi, R.; Brickner, M.; Chmielewski, J. Inhibition of the assembly of protein–protein interfaces. *Curr. Opin. Chem. Biol.* **1998**, *2*, 62.
38. Heldin, C.-H. Dimerization of cell surface receptors in signal transduction. *Cell* **1995**, *80*, 213.
39. Lemmon, M. A.; Schlessinger, J. Regulation of signal transduction and signal diversity by receptor oligomerization. *Trends Biol. Sci.* **1994**, *19*, 459.
40. Stahl, N.; Yancopoulos, G. D. The alphas, betas, and kinases of cytokine receptor complexes. *Cell* **1993**, *74*, 587.
41. Austin, D. J.; Crabtree, G. R.; Schreiber, S. L. Proximity versus allostery: the role of regulated protein dimerization in biology. *Chem. Biol.* **1994**, *1*, 131. Belshaw, P. J.; Ho, S. N.;

Crabtree, G. R.; Schreiber, S. L. Controlling protein association and subcellular localization with a synthetic ligand that induces heterodimerization of proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4604. For a recent disclosure, see: Koide, K.; Finkelstein, J. M.; Ball, Z.; Verdine, G. L. A synthetic library of cell permeable molecules. *J. Am. Chem. Soc.* **2001**, *123*, 398, and references cited therein.

42. Ihle, J. N.; Witthuhn, B. A.; Quelle, F. W.; Yamamoto, K.; Thierfelder, W. E.; Kreider, B.; Silvennoinen, O. Signaling by the cytokine receptor superfamily: JAKs and STATs. *TIBS* **1994**, *19*, 222.

43. Lamb, P.; Seidel, H. M.; Stein, R. B.; Rosen, J. The role of JAKs and STATs in transcriptional regulation by cytokines. *Annu. Rep. Med. Chem.* **1996**, *31*, 269.

44. Brivanlou, A. H.; Darnell, J. E., Jr. Signal transduction and the control of gene expression. *Science* **2002**, *295*, 813.

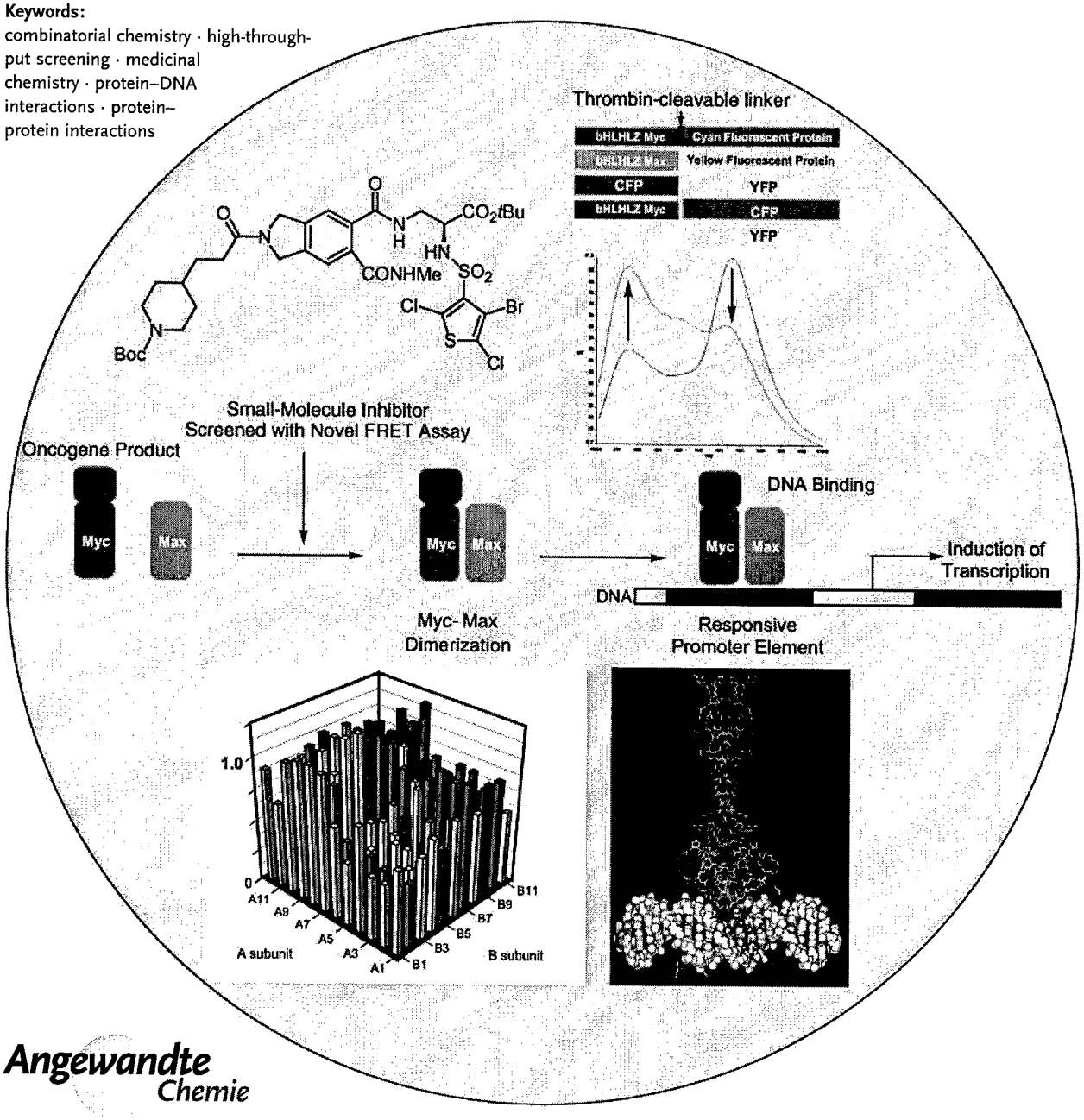
45. Seed, B. Making agonists of antagonists. *Chem. Biol.* **1994**, *1*, 125.

Combinatorial Chemistry

Solution-Phase Combinatorial Libraries: Modulating Cellular Signaling by Targeting Protein–Protein or Protein–DNA Interactions

Dale L. Boger, Joel Desharnais, and Kevin Capps*

Keywords:
combinatorial chemistry · high-throughput screening · medicinal chemistry · protein–DNA interactions · protein–protein interactions



  DOI: 10.1002/anie.200300574

The high-throughput synthesis and screening of compound libraries hold tremendous promise for drug discovery and powerful methods for both solid-phase and solution-phase library preparation have been introduced. The question of which approach (solution-phase versus solid-phase) is best for the preparation of chemical libraries has been replaced by which approach is most appropriate for a particular target or screen. Herein we highlight distinctions in the two approaches that might serve as useful considerations at the onset of new programs. This is followed by a more personal account of our own focus on solution-phase techniques for the preparation of libraries designed to modulate cellular signaling by targeting protein–protein or protein–DNA interactions. The screening of our libraries against a prototypical set of extracellular and intracellular targets, using a wide range of assay formats, provided the first small-molecule modulators of the protein–protein interactions studied, and a generalized approach for conducting such studies.

From the Contents

1. Introduction

Combinatorial chemistry has undergone rapid development and has provided an important approach for target validation and drug discovery. As a consequence of its extension from peptide and oligonucleotide synthesis, the majority of approaches have relied on solid-phase synthesis techniques and many reviews have treated this emerging science.[1–16] A complement to adapting solution-phase chemistry to polymer-supported combinatorial synthesis is the development of protocols for solution-phase combinatorial synthesis.[17–31] A number of approaches have been introduced for this purpose including dynamic combinatorial chemistry,[32,33] soluble polymer-supported combinatorial synthesis (pioneered by Janda et al.),[34–38] fluorous-phase synthesis (FPS)[39,40] and other tagging strategies,[41] precipitation techniques,[42–45] and resin-bound,[46–52] soluble polymer-bound,[53] and fluorous-phase bound[30,40] catalysts, reagents, or scavenging reagents.[46–52] Preceding this work, we introduced a simple approach that permits the multistep solution-phase synthesis of chemical libraries employing liquid–liquid and liquid–solid (ion-exchange resin) extractions to remove residual starting materials, reagents, and reagent byproducts providing the purified product (>95% pure) irrespective of the reaction efficiency.[57–70] It has been implemented in formats for the parallel synthesis of individual compounds[62,68] (1000-member libraries), modest-sized libraries composed of small mixtures (1000–10000-member libraries, 10–50 compounds per mixture) including positional-scanning and deletion-synthesis libraries,[63–65] or combinatorially assembled to provide large libraries (25000–1000000-member libraries, 10000–28000 compounds per mixture).[65] This flexibility allows the method to be adopted in a format compatible with any screening objective. Thus, its implementation is convenient for either lead discovery or lead optimization and produces the library members on a scale that allows their repeated use in screening without resynthesis. Such libraries have been prepared in which members have been found to interfere with (antagonists) or mimic (agonists) extracellular or intracellular protein–protein interactions, inhibit intracellular enzymes, or modulate intranuclear protein–DNA interactions. The chemistry has proven applicable to natural product scaffolds, and cyclic (depsi)peptides and possesses a scope that exceeds what one might initially imagine based on its simplicity.[57–99]

This technology is being developed as an integral component of efforts to modulate cellular signaling by inhibiting, promoting, or mimicking protein–protein or protein–DNA interactions, especially those that lack existing small-molecule leads and target structural information. Despite the prevailing bias that such targets might not prove viable for small-molecule intervention,[100,101] the screening of our libraries provided the first small-molecule modulators of the protein–protein interactions being studied, validated the targets for therapeutic intervention, and provided a generalized approach to conducting such investigations.[102] Herein, we review these efforts by first detailing methodology introduced for the solution-phase synthesis of chemical libraries followed by a more personal summary of our applications in the validation of a series of new therapeutic targets.

2. Solution-Phase versus Solid-Phase Combinatorial Chemistry

2.1. Key Distinctions of Solution-Phase Combinatorial Approaches

Solid-phase combinatorial chemistry has been widely implemented for lead discovery and lead optimization

[*] Prof. D. L. Boger, J. Desharnais, K. Capps
Department of Chemistry and
The Skaggs Institute for Chemical Biology
The Scripps Research Institute
10550 North Torrey Pines Road, La Jolla, CA 92037 (USA)
Fax: (+1) 858-784-7550
E-mail: boger@scripps.edu

 DOI: 10.1002/anie.200300574

because of the simplicity by which the compounds can be isolated once they are attached to solid supports. Several advances in the implementation of solid-phase library synthesis paved the way for its widespread acceptance including the spatially arrayed multipin peptide synthesis of Geysen et al.,[103,104] the tea-bag method of peptide synthesis introduced by Houghten et al.,[105] phage display from Smith et al.,[106,107] the spot or disc synthesis developed by Frank et al.,[108,109] and especially the (portion-mixing) split-and-mix solid-phase synthesis on beads introduced by Furka et al.[110–112] and also disclosed by Houghten et al.[113] (divide, couple, and recombine), and Lam et al.[114–116] (split synthesis), and Affymax's light-directed, spatially addressable, immobilized parallel synthesis.[117] Further promoting the widespread acceptance of split-and-mix solid-phase synthesis, several identification techniques were introduced including iterative[110,118,119] and recursive[120] deconvolution, or nucleotide-,[121,222] peptide-,[123–125] chemical-,[126–128] radiofrequency-,[129,130] color-,[131] and shape-encoded[132] solid-supported libraries. Similarly, advances in microarray synthesis continue to improve large-scale spatially arrayed parallel synthesis on a variety of solid supports.[133–136] In addition to the many merits of solid-phase synthesis, the isolation of the immobilized product by simple filtration permits the use of large reagent excesses to effect high-yield conversions for each of the steps (Table 1). However, its scale can be restricted by the amount of required solid support and its loading capacity and the production of significant quantities of each library member can be cumbersome and expensive for large libraries. Its use requires functionalized substrates and solid supports for attachment, compatible spacer linkers, orthogonal attachment and detachment techniques often with the release of spectator functional groups, specialized methods for monitoring the individual steps of a multistep synthesis often including orthogonal capping strategies for blocking residual substrate, and does not permit the purification of resin-bound intermediates. This last feature necessarily produces the released product of a multistep sequence in an impure state and requires that each step of a multistep synthesis on each substrate proceed with high efficiency. Advances addressing each of these limitations have been made which continue to improve the implementation of solid-phase synthesis. Some of the most notable introductions include traceless linkers,[137–139] the resin release of product only upon successful synthesis,[140,141] resin-capture of only the desired reaction products,[142] the introduction of safety-catch linker[143–145] and backbone amide linker[146,147] strategies, multidirectional resin-cleavage methods,[148,149] and improved resin properties and loading capacities.[34–38] Limitations that cannot be altered are that solid-phase synthesis is incompatible with the use of heterogeneous catalysts or reagents, is necessarily restricted to a linear (as opposed to convergent) synthetic strategy, is not easily amenable to the synthesis of mixture libraries, and is not amenable to selection or dynamic library screening[4] involving target-assisted synthesis.[150–154] These latter limitations are discussed in more detail in the following sections.

Table 1: Advantages (+) and disadvantages (−) of solution-phase and solid-phase combinatorial synthesis.

	Solid-phase		Solution-phase
+	Simple removal of excess reagents and reactants	−	Removal of excess reagents and reactants limits scope
+	Automation straightforward	+	Automation by liquid–liquid techniques
+	Split and mix synthesis simplifies preparation of large libraries	+	Mixture synthesis simplifies preparation of large libraries
+	Pseudo–dilution effects	+	Chemistry not limited by support
−	Adapt chemistry to solid phase and develop linking/cleavage strategies	+	Avoids extra steps for linking/cleavage
−	Reaction monitoring difficult	+	Monitor by traditional techniques
−	No purification possible	+	Purification possible after each step
−	Linear, but not convergent synthesis	+	Convergent or linear synthesis
−	Limited scale	+	Unlimited amounts (scales) available
−	Cannot conduct mixture synthesis	+	Mixture or parallel synthesis

At the time we initiated our efforts, no reports of the extension of the solid-supported synthesis of chemical libraries to solution-phase libraries had been described and a series of related studies including those of Smith et al.,[155] Rebek, et al.,[156–159] Pirrung et al.,[160] and Nielsen et al.,[161] and the popularization of multicomponent solution-phase

Dale L. Boger received his B.Sc. from the University of Kansas (1975) and Ph.D. degree from Harvard University (1980). He returned to the University of Kansas (Department of Medicinal Chemistry) then in 1985 moved to Purdue University (Department of Chemistry). Since 1991 he is the Richard and Alice Cramer Professor of Chemistry at The Scripps Research Institute. His research interests include synthetic methods, the total synthesis of natural products, heterocyclic, bioorganic, medicinal chemistry, and combinatorial chemistry, DNA–agent and protein–ligand interactions, and antitumor agents.

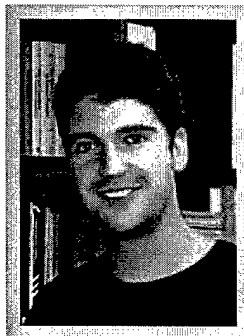
Joel Desharnais received a B.Sc. degree in Chemistry from the University of Montreal (1995) and his Ph.D. degree in Chemistry from the University of Montreal (2000) working with Professor James. D. Wuest in the field of supramolecular chemistry. Following this he was a postdoctoral fellow at The Scripps Research Institute working with Professor Boger, where his research interests entailed the synthesis of small-molecule agonists/antagonists of protein–protein interactions using solution-phase combinatorial chemistry and the synthesis of new folate-based enzyme inhibitors. He joined Acidophil (San Diego) in 2003.

reactions by Ugi et al.[162] and Armstrong et al.[163] were disclosed around the time of our initial contributions. The distinction of our work being that it constituted a multistep, not a single-step, strategy for the solution-phase synthesis of chemical libraries.

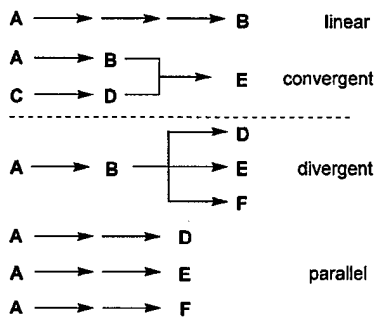
Given that both solution-phase and solid-phase sample manipulation are easily automated, the major limitation to the solution-phase synthesis of chemical libraries is the isolation or purification[164] of the library members. If the advantages of sample isolation characteristic of solid-phase synthesis can be incorporated into a solution-phase synthesis, its less limiting scale, expanded repertoire of chemical reactions, direct production of soluble intermediates and final products for purification or assay, and the lack of required linking, attachment and detachment, or capping steps make solution-phase library synthesis especially attractive (Table 1). It is the only approach readily amenable to convergent synthetic strategies, the synthesis of mixture libraries, or use of dynamic libraries. Consequently, we considered a number of techniques that are available for simplifying the isolation of solution-phase samples and focused on one of the most attractive involving isolation and purification by liquid–liquid or solid–liquid (ion-exchange resins) acid–base extractions. Descriptions of the initial libraries generated and their subsequent extensions, which included the introduction of solid-supported reagents,[57–63] solid-supported workup reagents (scavengers,[165] ion-exchange resins[57–59]) including the introduction of immobilized scavengers and reagents conducted within CombiChem in collaboration with L. Caporale,[165] purification by size-exclusion chromatography filtration,[62,63] and the generation of non-amide libraries have been detailed.[71]

2.2. Synthetic Strategies: Linear, Convergent, and Divergent Synthesis

Traditional organic synthesis has provided two widely recognized strategies for synthetic design: linear versus convergent synthesis (Scheme 1). The merits of a convergent synthesis are widely recognized; it typically provides a shorter (number of steps), more efficient (higher overall yield), and more effective (larger sample size) approach that is technically simpler (material balance) to execute than a linear



Kevin Capps graduated with a B.Sc. in Chemistry from the University of Maryland in 2000 while conducting undergraduate research in the laboratory of Professor Jeffery Davis. He is currently a graduate student at Scripps.



Scheme 1. Linear versus convergent and divergent versus parallel synthesis.

synthesis. Since solid-phase synthesis is limited to a linear synthetic strategy, it is not surprising that it did not find widespread use in traditional organic synthesis. In fact, the applications where solid-phase synthesis did find widespread acceptance prior to the advent of combinatorial chemistry (solid-phase peptide and oligonucleotide synthesis) are those ideally suited for linear synthesis involving the repetitive use of a single highly optimized reaction. A third, less widely recognized, synthetic strategy is the divergent synthesis where a class of compounds is prepared from a single common intermediate. Although widely used, the formalization (definition) of this strategy was first detailed along with a strategy for divergent aromatic annulations.[166–168] Today, the divergent introduction of diversity is the basis of most combinatorial chemistry strategies.[169] The advantages of such a late-stage divergent approach relative to the parallel synthesis of a series of related compounds are analogous to those of a convergent versus linear synthesis (more efficient, larger sample size, material balance).

The strategies are not mutually exclusive for diversity-oriented synthesis. Either a linear or convergent synthetic strategy can be employed (Figure 1). Solid-phase synthesis is necessarily limited to a linear synthetic strategy whereas solution-phase synthesis may be adapted for either a linear or

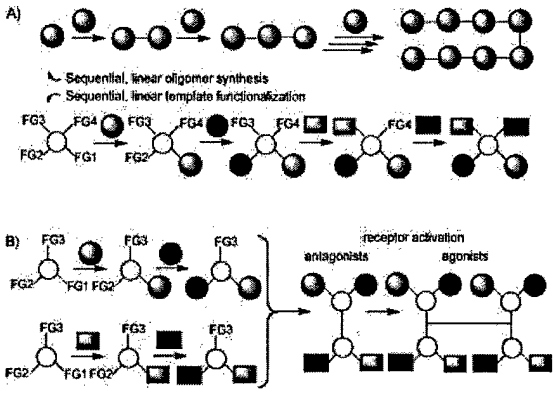
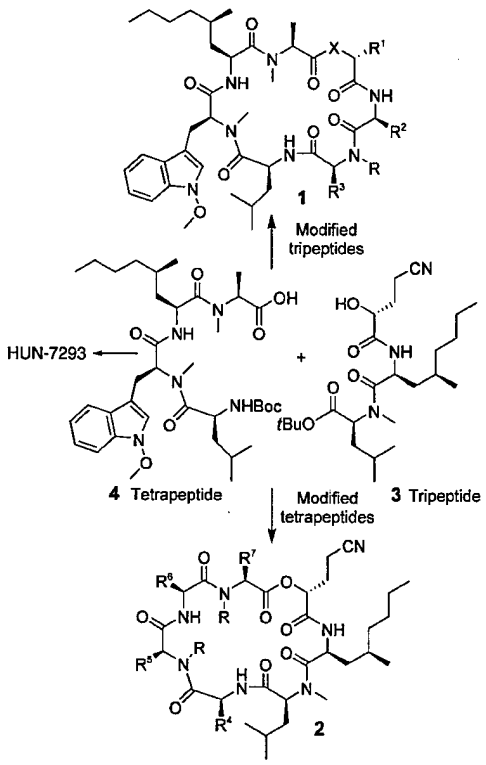


Figure 1. Convergent versus linear synthesis. A) Linear, divergent synthesis with multiplication of diversity (solid or solution phase), B) convergent synthesis with multiplication of diversity (solution phase only).

convergent synthetic strategy. Moreover, the convergent synthesis is especially well suited for solution-phase synthesis and would be precluded by solid-phase techniques where the combining components are on mutually exclusive solid phases.

From the start, an important element of our efforts to probe protein–protein or protein–DNA interactions was the ability to screen for binding agents for each macromolecular partner (screen for antagonists) which could then be linked to promote the protein–protein interaction in the absence of an endogenous ligand (e.g., agonists for receptor homo- and heterodimerization; Figure 1). This was the genesis of the use of the iminodiacetic acid template for the synthesis of libraries and it has provided the first generalizable approach for the discovery or development of agonists from antagonists.[61,63,67,73,170,171] This convergent mix-and-match dimerization linkage of macromolecular binders was effected by a variety of techniques including diamide formation, olefin metathesis,[60,61,64] palladium-catalyzed biaryl coupling,[62] and Stille acetylenic–aryl couplings[63] and can only be accommodated by solution-phase, not solid-phase techniques.

An additional example of the superimposition of a divergent, diversity-oriented synthesis onto a convergent synthetic strategy may be found in work that provided a library of HUN-7293 analogues (**1** and **2**, Scheme 2) which was used to explore each structural detail of the natural product.[76] As discussed in Section 6.1, a solid-phase synthesis of the library would require a parallel linear synthesis of each analogue, whereas the implementation of a solution-phase convergent synthesis allowed the single synthesis of the key tri- and tetrapeptide subunits of the natural product (**3** and **4**) and their respective combination with diversified tetra- and tripeptides containing single-point changes in the HUN-7293 structure (Scheme 2). Thus, the advantages commonly associated with a convergent synthesis were combined with those of a divergent synthesis to provide a family of closely related structures.



Scheme 2. Convergent synthetic strategy for the generation of a library of HUN-7293 analogues. Boc = *tert*-butoxycarbonyl X = O, NH.

2.3. Solid-Phase Split-and-Mix versus Solution-Phase Mixture Synthesis and Deconvolution (Identification) Methods

Libraries of modest size can be prepared easily by either parallel solid-phase or solution-phase synthesis, both of which can provide individual compounds. The advent of split-and-mix solid-phase synthesis and the related tea-bag methods which provide one-compound-per-bead libraries made large libraries accessible by solid-phase synthesis. The concurrent introduction of the identification (tagging) methods detailed in Section 2.1[121–132] greatly enhanced the value of such strategies by providing an immediate identification of screening hits. The most valuable of these encoding technologies being those that do not add complexity to the library synthesis (Table 2).

Mixture synthesis provides for solution-phase synthesis what split-and-mix synthesis provided for solid-phase techniques (Table 2).[172] Unlike solid-phase synthesis where the polymer-bound substrate must be the stoichiometrically limiting reaction partner, either the substrate or the reacting attachment groups may be limiting in solution-phase chemistry. This situation dictates the use of split-and-mix synthesis in

Table 2: Comparison of library techniques.

Technique	Single compound or mixture	Speed of synthesis	SAR[a] retrieval	Utility
parallel synthesis	single	slow	fast	lead optimization
mixture synthesis (scanning/deletion deconvolution)	mixture	fast (fast)	slow (fast)	lead identification
parallel, spacial array	single	moderate	fast	lead identification or optimization
split and mix	mixture (one compound per bead)	moderate	slow	lead identification or optimization
encoded split and mix	mixture (one compound per bead)	moderate	moderate	lead identification or optimization
mix and sort (microreactors)	single	moderate	fast	lead identification or optimization

[a] SAR = structure activity relationship.

the solid-phase synthesis to accommodate different reaction rates,[173] whereas the simpler method of mixture synthesis with limiting attachment-group stoichiometry may be used in solution to ensure all library members are generated in equal amounts. The cases where mixture synthesis has been conducted on the solid-phase enlists an excess of the reacting attachment groups in adjusted concentrations to accommodate the different reaction rates and requires that this rate information be available at the onset of the mixture synthesis. In contrast, the implementation of solution-phase mixture synthesis only requires the ability to remove residual substrate. Although this is not possible with solid-phase synthesis, it is readily accomplished in our efforts by simple acid–base liquid–liquid or liquid–solid extraction.

Analogous to encoding strategies for solid-phase split-and-mix synthesis of large libraries, the two important and complementary techniques of positional scanning[173] and deletion synthesis[64,65] permit the immediate identification of active leads from large mixture libraries (Figure 2). Because the preparation of positional-scanning and deletion-synthesis libraries entails implementing mixture synthesis, they are not easily adapted to solid-phase techniques. Typically, global observations are detected with a positional-scanning library and useful lead structures (although not necessarily the most potent member)[79,86,175] with defined properties are identified. Deconvolution with positional-scanning libraries is more sensitive because an increase in activity is measured, and this allows the identification of less-active lead structures. Deletion synthesis, on the other hand, screens for a loss of activity, provides less global information, but is better at identifying a uniquely potent library member.[64,65,174] The combination is more powerful than either technique alone.[64] Importantly, the lead identification is deduced from a single round of screening. Thus, not only are positional-scanning or deletion-synthesis libraries much less time intensive to prepare than the parallel synthesis of individual compounds or small mixtures and technically less demanding than spacially arrayed or encoded split-and-mix library synthesis, but they produce depository libraries for use in multiple screens capable of immediate deconvolution in a single round of testing. Although a subtle distinction and not typically a consideration, the convergent synthesis of dimeric libraries providing symmetrical and unsymmetrical combinations cannot be carried out with positional scanning libraries, but must use deletion synthesis to identify active unsymmetrical dimers.[64] Relative to the screening of individual compounds or small mixtures, the more subtle structure–activity relationship information contained in a library is often missed. However, the disadvantages associated with the loss of this information is balanced against the ease of synthesis of the parent libraries and must be judged in light of the screening objectives (lead identification or lead optimization), and the nature of the target.

A number of analytical techniques have also emerged as useful approaches to lead identification from complex mixtures, and these include assay (detection) and identification by mass spectrometry,[176] NMR spectroscopy,[177,178] and capillary electrophoresis,[179] and have expanded the utility and dynamic range for screening such mixture libraries.

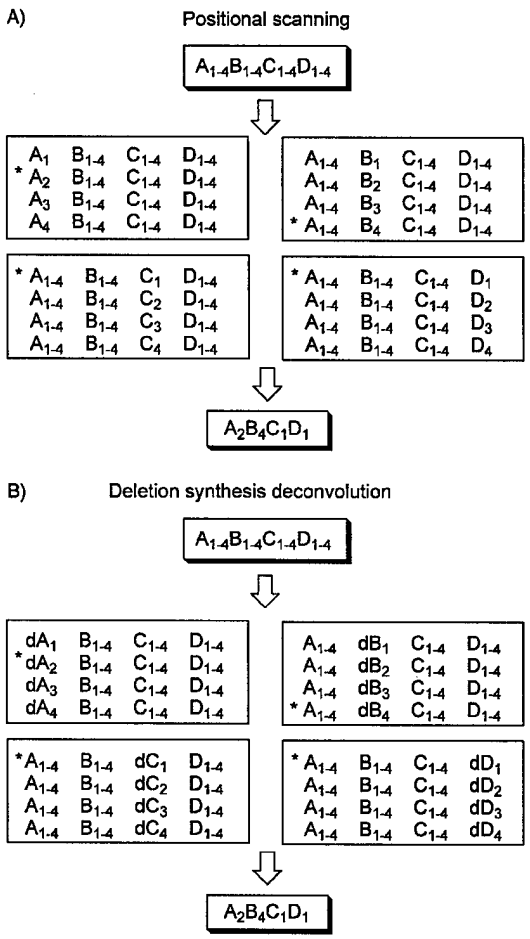
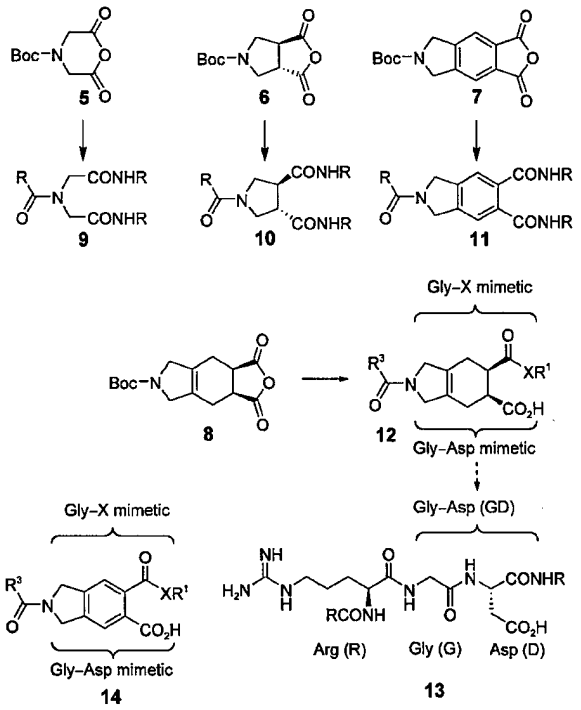


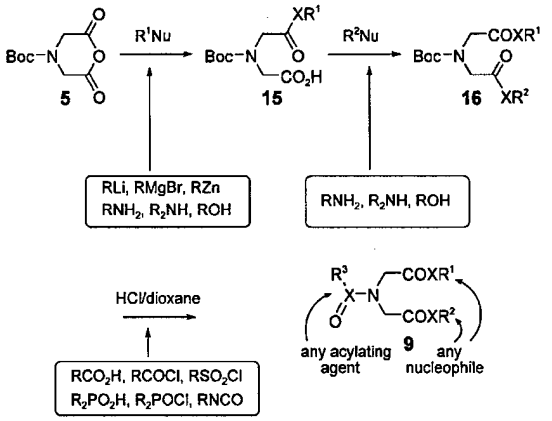
Figure 2. Examples of positional-scanning and deletion-synthesis deconvolution techniques for a library that contains four variable positions with four substitutions at each position. Positional scanning: In each sublibrary, an individual component is placed in one position but mixtures are used in all other positions. The most active sublibrary in each set (*) indicates the optimal substituent for that position. Deletion synthesis: In each sublibrary, an individual component is missing in one position (e.g. dA_2 signifies all A substituents except A_2), but full mixtures are used in all other positions. The least active sublibrary in each set (*) indicates the optimal substituent for that position, that is, the missing component is the most active.

3. Solution-Phase Library Preparation: Cyclic Anhydride Templates

A summary of all the work on solution-phase libraries is beyond the scope of this article and excellent coverage is provided in a number of recent reviews.[17–31] However, we would like to briefly describe the basis of our own library preparations highlighting the design elements considered in the context of the biological targets we addressed. Template **5** (Scheme 3), the first of a representative set of five- and six-membered cyclic anhydride-based templates that have been examined, consists of a densely functionalized core which imposes little structural or conformational bias that might



Scheme 3. Representative templates.

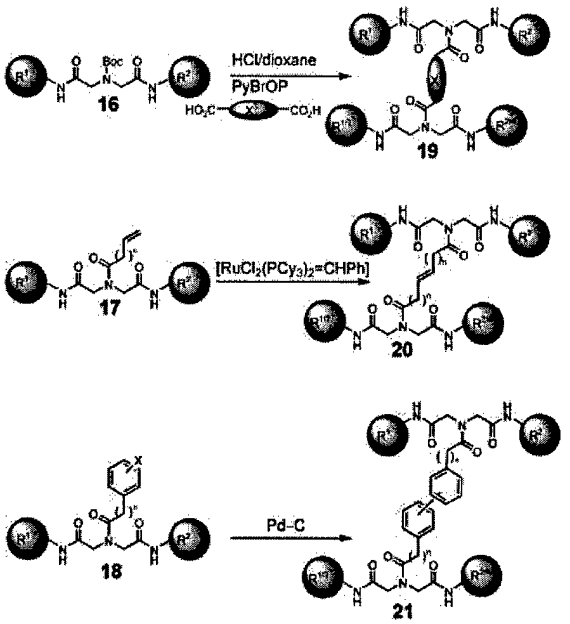


Scheme 4. Three-step solution-phase synthesis of iminodiacetic acid chemical libraries. Compound **15** is purified by washing with acid, **16** and **9** by washing with acid and base. (The R groups of the reagents in boxes become R^1, R^2, R^3.)

limit its use.[57,58] Interesting examples of alternative templates include the dipeptidomimetic template **8**, which contains a rigid bicyclic core with a plane of symmetry that allows it to function as a Gly–Asp or more general Gly–X mimetic.[57,59] Its symmetrical structure, like that of **5**, contains three positions that can be sequentially functionalized enabling the synthesis of libraries with up to three variable groups. A more rigid variant of a Gly–X mimetic **14**, isoindoline-5,6-dicarboxylic acid (**11**), has also been examined.[69] Importantly, an array of such templates, such as **5–7** (Scheme 3), displays the introduced diversity in conformationally constrained or different spacially defined locations providing rich structure–activity relationship information on the target of interest in a single round of testing.

The anhydride template is activated for the first functionalization, which is conducted by nucleophilic addition, and generates a free carboxylic acid group as its second functionalization site (Scheme 4). Although a subtle design element, this release of the second functionalization site concurrent with the first diversification not only provides an isolation and purification handle, but it avoids a separate deprotection step for the subsequent diversification. In addition, the use of symmetrical anhydrides insures only a single regioisomer of the product is generated simplifying the potential complexity of the library. The final diversification entails deprotection of the secondary amine which unmasks a third coupling position. At each step, the released functionality aids not only the isolation, but provides a method for separation of each of the intermediates and final products from starting materials, reactants, reagents, and their byproducts by simple liquid–liquid or liquid–solid acid–base extraction providing pure materials (typically >95%) independent of reaction efficiencies, and a number of classes of nucleophiles and electrophiles can be used to functionalize the template (Scheme 4).[71]

This approach has been used to provide libraries in a variety of formats (individual compounds, small mixtures, large mixtures, positional-scanning, deletion-synthesis libraries) applicable to both lead generation or lead optimization. In fact, it was in this context that deletion-synthesis libraries[64] were introduced and compared with other library formats including positional scanning.[65,79,86] Although this work is most often cited for its isolation and purification by liquid–liquid acid–base extraction, the initial disclosures[57–59,68] detailed one of the first applications of polymer-bound immobilized reagents in solution-phase library synthesis (polymer supported EDCI (PS-EDCI), EDCI = 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide) and also described the first use of ion-exchange resins in liquid–solid "extraction" workups.[58] The approach has been shown to dependably deliver pure compounds in 5–150 mg quantities (>95% pure) and multiple libraries have been assembled providing over 40000 compounds to date. It avoids the disadvantages of solid-supported synthesis, including the more restrictive scale, the required functionalized substrates and supports, compatible linkers, and the requirements for orthogonal attachment and detachment chemistry. It does not require specialized methods for monitoring each step of multistep syntheses, allows the purification of intermediates, and provides the final pure products directly for use in binding or functional assays. Extensions of these studies for generating symmetrical or unsymmetrical chemical libraries suitable for probing receptor and protein homo- and heterodimerization events were described (Scheme 5). The dimerization linkage was accomplished not only through amide-bond formation with dicarboxylic acids, but also by Stille coupling,[63] by the first application of an intermolecular

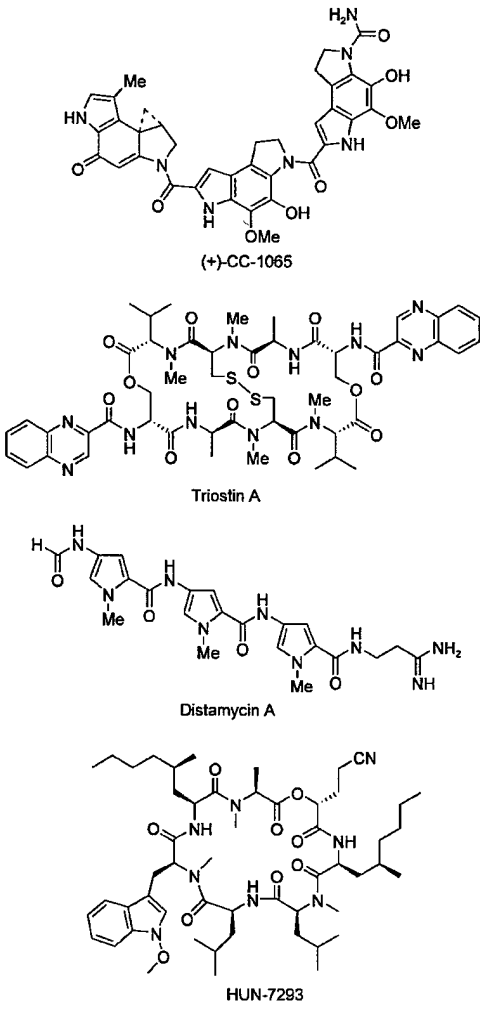


Scheme 5. Homo- or heterodimerization of libraries for protein–protein dimerization. The deconvolution of the unsymmetrical dimerization products is by scanning or deletion synthesis. Cy = cyclohexyl, PyBrOP = bromotris(pyrrolidino)phosphonium hexafluorophosphate.

olefin-metathesis reaction in the generation of chemical libraries,[60,61,64] and by one of the first disclosed applications of solid-supported catalysts (Pd-C) in solution-phase library preparation that was used to promote a biaryl coupling (see Scheme 5).[62] In this context, the use of size exclusion chromatography (filtration) for the rapid purification of such dimer, or higher oligomer, libraries was also demonstrated.[62]

Thus, the preparation and dimerization linkage of iminodiacetic acid diamides can be conducted in a reaction sequence that requires only three steps. In addition to the multiplication of the diversity that arises through the combinatorial dimerization linkage of the iminodiacetic acid diamides, the solution-phase synthesis of the intermediates permits their direct linkage which would be precluded by solid-phase synthesis techniques. As such, the strategy is uniquely suited for taking advantage of such dimerization (convergent) strategies utilizing a limited number of synthetic steps. This modular approach to the generation of libraries is especially well-suited for the discovery of antagonists or agonists of receptor and protein homo- and heterodimerization. Simple binders (i.e., **16–18**, Scheme 5) can serve as antagonists of ligand-induced receptor or protein dimerization. Covalently linked symmetrical dimers (i.e., **19–21**) can be utilized to promote receptor or protein homodimerization whereas unsymmetrical dimers can be utilized to promote receptor or protein heterodimerization. Thus, both antagonists and agonists may be developed depending on the therapeutic application and this was a major factor in our adaptation of this approach. Notably, the size and characteristics (cell permeability) of compounds targeting intracellular macromolecules may place restrictions on the library composition, whereas libraries of compounds addressing extracellular targets including cell surface receptors and proteins may not be influenced by such limitations.

One might arguably suggest from an examination of the anhydride templates that the structural complexity, available diversity, and drug-like properties of the library members, or scope of the technology may be restricted producing libraries of limited value. However, the present 40000-member library has been remarkably successful in providing leads for essentially any target it has been screened against. Counter to the prevailing trend of preparing libraries of increasing molecular rigidity and complexity, it may well be that this success can be attributed to the library's appropriate "intermediate" complexity where each library member embodies multiple pharmacophores more suited for lead discovery.[180] It should also be emphasized that alternatives to the amide couplings even for the initial library generations have been disclosed.[71] In addition and although this is discussed in



Scheme 6. Natural product templates where analogues were synthesized using the solution-phase combinatorial approach.

greater detail in Sections 5 and 6, complex natural product libraries (Scheme 6) have been assembled using this technology indicating that the scope extends well beyond what one might initially suspect.

4. Targeting Protein–Protein Interactions

As attractive as the potential is for modulating aberrant cellular signaling[181–190] by inhibiting, promoting, or mimicking protein–protein[100,101,181,182] or protein–DNA interactions, the general consensus has been that such targets represent stubborn problems for small-molecule therapeutic intervention. The challenges posed by such targets arise from the large surface area generally covered by the two interacting proteins (approximately 800 Å^2 per protein on average), the relatively flat interface between binding proteins, and the range of affinities between proteins depending on their physiological function.[181] Although there are many examples of antibodies, dominant negative proteins, and medium-sized peptides that inhibit protein–protein interactions, the disclosure of small molecules has not been as extensive. Because the nature of protein–protein interactions differ widely (hydrophobic versus hydrophilic, the surfaces can be relatively flat and/or have buried crevices), since the affinities span a considerable range (K_d = 10^{-6}–10^{-12} M), and the response to inhibition may vary considerably (the need to simply perturb the equilibrium versus need for complete inhibition), it is fair to assume that some protein–protein interactions may prove to be better targets than others for small-molecule intervention. Table 3[191–225] summarizes many of the recent promising results of investigations on a range of targets and Scheme 7 shows the structures of some of the small-molecule modulators. It is clear that a reevaluation of the prevailing bias against such pro-

Table 3: Small molecule modulators of protein–protein interactions.

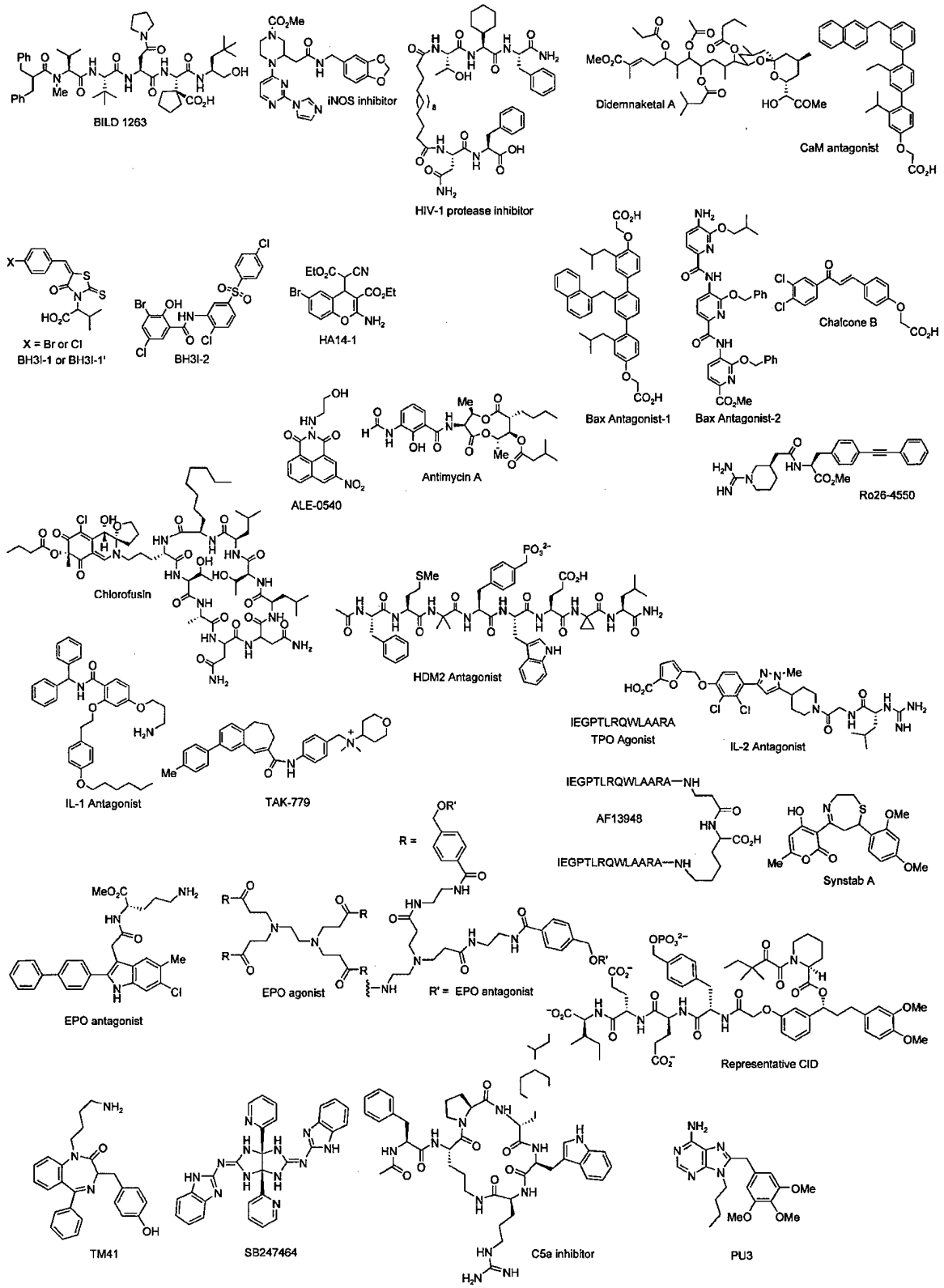
Protein–protein Interaction	Compound Source	Screen	Lead Structure[a]	Activity	Ref.
Ribonucleotide Reductase (dimerization)	Peptide analogue	Competitive binding assay	BILD 1263	Antagonist IC_{50} = 0.3 nM	[197]
iNOS (dimerization)	Encoded solid-phase library	Functional assay	iNOS inhibitor	IC_{50} = 0.6 nM	[198]
HIV Protease (dimerization)	Designed inhibitor	Enzyme inhibition	HIV-1 protease inhibitor	K_i = 310 nM	[199a]
HIV Protease (dimerization)	Natural product	Enzyme inhibition	Didemnaketal A	IC_{50} = 2 μM	[199b]
HIV-1 Integrase (dimerization)	Interface peptides	Enzyme inhibition	–	IC_{50} = 2–4 μM	[200]
CaM/smMLCK	α-helix mimetic	Enzyme inhibition	CaM antagonist	IC_{50} = 9 nM	[201]
MMP-2/$\alpha_v\beta_3$	Solution phase library	Radiolabeled binding assay	**27** (Figure 11)	Angiogenesis inhibitor	[77,78]
Paxillin/α4	Positional scanning library	ELISA	**A7B7C7** (Figure 13)	Antagonist IC_{50} = 300 nM	[79]
Bax/Bcl-x_L	16320 compound library	Fluorescence polarization assay	BH3I-1 BH3I-1′ BH3I-2	Antagonist IC_{50} = 8–16 μM	[202]
Bax/Bcl-x_L	Natural product	Functional assay	Antimycin	Antagonist	[203]
Bax/Bcl-x_L	Virtual screening of 193383 compounds	Fluorescence polarization assay	HA14-1	Antagonist IC_{50} = 9 μM	[204]
Bax/Bcl-x_L	Designed α-helix mimetic	Fluorescence polarization assay	Bax antagonist-1	K_D = 114 nM	[205a]
Bax/Bcl-x_L	Designed α-helix mimetic	Fluorescence polarization assay	Bax antagonist-2	K_i = 1.6 μM	[205b]
MDM2/p53	Natural product	ELISA	Chlorofusin	Antagonist IC_{50} = 4.6 μM	[206a]
MDM2/p53	Unknown	ELISA, NMR	Chalcone B Antagonist	IC_{50} = 49 μM K_D = 90 μM	[206b]
HDM2/p53	Designed inhibitor	ELISA	HDM2 Antagonist	IC_{50} = 5 nM	[207]
Myc/Max	Solution phase library	FRET	**28** (Figure 17)	Antagonist IC_{50} = 20 μM	[80]
LEF-1/β-catenin	Solution phase library	Reporter assay	**A1B2C8** (Figure 18)	Antagonist	[71]
NGF/p75 NGF/TrkA	Commercial libraries	Radiolabeled binding assay	ALE-0540	IC_{50} = 6 μM IC_{50} = 4 μM	[208]
HIV-1 (RANTES)/CCR5	Corporate library	Radiolabeled binding assay (RANTES)	TAK-779	Antagonist IC_{50} = 1.4 nM	[209]
IL-2/IL-2R	Designed inhibitor	Radiolabeled binding assay	Ro26-4550	Antagonist IC_{50} = 3 μM	[210a]
IL-2/IL-2R	Structure-based design + fragment assembly	Competitive binding	IL-2 Antagonist	IC_{50} = 60 nM, K_d = 100 nM	[210b]
IL-1/IL-1R	Designed inhibitor	Scintillation proximity assay	IL-1 Antagonist	IC_{50} < 10 μM	[211]
EPO/EPOR	Phage display	Radiolabeled binding assay	EMP1 (Figure 20)	EPO Agonist IC_{50} = 0.2 μM	[319–321]
EPO/EPOR	Corporate library	Radiolabeled binding assay	EPO antagonist	IC_{50} = 60 μM	[330]
	Modified to	Radiolabeled binding assay	EPO agonist	EPO Agonist IC_{50} = 4.4 μM	[330]
EPO/EPOR	Solution phase library	Radiolabeled binding assay	**30** (Figure 19)	EPO Agonist EC_{50} = 8 μM	[82]
TPO/TPOR	Phage display	ELISA	TPO Agonist	EC_{50} = 400 nM	[212]
	Modified to	ELISA	AF13948 (dimer)	EC_{50} = 100 pM	[212]

Table 3: (Continued)

Protein–protein Interaction	Compound Source	Screen	Lead Structure[a]	Activity	Ref.
TPO/TPOR	Corporate library	Radiolabeled binding assay	TM41	TPO Agonist	[213]
GCSF/GCSFR	Corporate library	Functional assay	SB247464	Agonist	[214]
C5a/C5aR	Design inhibitor	Radiolabeled binding assay	C5a antagonist	$IC_{50}=300$ nM	[215]
FKBP/Cn	Natural product	–	FK506	Cn inhibitor	[216]
Cyclophilin/Cn	Natural product	–	Cyclosporin A	Cn inhibitor	[216217]
FKBP/FRAP	Natural product	–	Rapamycin	Immunosuppressant	[218]
Chemical inducers of dimerization (CIDs)	Synthetic constructs	Various	Representative CID	Various	[192–196]
Ribosome (RNA)/protein	Natural products	Various	Macrolides Tetracycline Oxazolidinones Streptogramin Chloramphenicol TAN 1057 Pestalone Cycloheximide Hygromycin Blasticidin Virginamycin Sparsomycin Puromycin Clindamycin	Antibiotics	[219]
Hsp90/Client Protein	Natural products	Various	Geldanamycin Herbimycin Radidicol 17-AAG	Antitumor compounds	[220a]
Hsp90/Client Protein	Designed ligand	Competitive binding	PU3	$EC_{50}=15$–20 μM	[220b]
α/β tubulin	Natural products	Various	Colchicines Vincristine Taxol Epothilones	Mitotic inhibitors	[221a]
α/β tubulin	Commercial Library	Cell-based screen	Synstab A	Mitotic Inhibitors	[221b]
Topoisomerase/DNA	Natural product	Various	Camptothecin	Antitumor compounds	[222]
PG/Transpeptidase	Natural product	Various	Vancomycin	Antibiotic	[223a]
LI II/Transglycosylase	Natural product	Various	Ramoplanin	Antibiotic	[,223b]
Nuclear hormone receptors	Various	Various	Various	Agonists and antagonists	[224]
RGD mimetics	Various	Various	Various	Antagonists	[240–252]
SH2 domains	–	–	–	–	[181,225a]
SH3 domains	–	–	–	–	[225b]
WW domains	–	–	–	–	[225c]
PDZ domains	–	–	–	–	[225d]

[a] Lead structure is shown in Scheme 7 unless indicated otherwise.

tein–protein interaction targets is warrented.[100–102,171,181,182] Interestingly, fewer of these examples represent results derived from screening combinatorial libraries than might be expected and most that do relied on phage display to generate initial peptide leads. The examples in Table 3 represent a wide range of strategies despite the commonality of all in perturbing a protein–protein interaction. For example, those that entail the inhibition of enzyme action do so by a range of mechanisms including directly preventing their homodimerization (HIV protease) or heterodimerization (ribonucleotide reductase), or inducing a conformational change that precludes dimerization (iNOS), effecting their localization (MMP-2, CaM), access to the substrate (transpeptidase, transglycosylase), stabilization of an intermediate ternary complex (topoisomerase), or promoting inhibitory ternary complexes with additional proteins (calcineurin (Cn)). Not included in this summary are enzyme active-site inhibitors that block a peptide or protein substrate binding to the enzyme, small-molecule agonists and antagonists of G-protein-coupled receptors, and all the cell surface integrin inhibitors based on the RGD recognition motif. The G-protein-coupled receptor antagonists typically bind the receptor transmembrane domain and do not directly (perhaps allosterically) perturb a receptor–ligand interaction. In contrast, the cell surface integrin inhibitors have been the subject of several reviews[240–252] and constitute a beautiful example of the discovery of a simple, small, and general protein–protein interaction motif that lends itself to subsequent antagonist development. Similarly, Src-homology 2 domains (SH2), SH3, WW, and PDZ domains constitute common protein–protein binding sites with small, focused interfaces for protein localization or signal transduction that have been the subject of small-molecule screening or design efforts.[101,181,225] In addition, there are several historically important modulators of protein—protein interactions that have defined in such discussions important therapeutic targets. Although these are often neglected, they are superb examples of achieving productive therapeutic endpoints by interfering with a protein–protein (or protein–DNA) interaction. Notably, there may be significant distinctions in the work, design of the library, screening methods, and challenge of the investigations depending on whether it is an antagonist or agonist of a protein–protein interaction that is sought. To date, the consensus is that the discovery of antagonists may be much



Scheme 7. Representative small-molecule modulators of protein–protein interactions listed in Table 3.

easier than an agonist.[73,170,171] As such, the efforts related to the discovery of EPO, TPO, and GCSF agonists (mimetics; Table 3) constitute remarkable achievements. Finally, the development of general methods for establishing the selectivity of such antagonists or agonists is a challenge that still has not been faced by the field and that may constitute a major reason for the reluctance to address protein–protein interactions as useful therapeutic targets.

Several very clever ways of enlisting chemically prepared bifunctional protein binders that induce dimerization to achieve a biological endpoint or objective have been detailed.[191] These chemical inducers of protein dimerization (CIDs) pioneered by Schreiber and Crabtree et al. have been used to modulate cell membrane receptor signaling,[192] to control gene expression,[192c,193] to selectively antagonize cellular processes,[194] to identify protein–ligand partners,[195] and even to target protein heterodimers with no established ligand.[196]

Cell growth, differentiation, migration, and apoptosis are regulated in part by growth factors or cytokines. These factors are unable to penetrate the cell membrane and exert their effects by binding to cell surface receptors. In many instances, such receptors are activated by ligand-induced dimerization or oligomerization.[183–186] In addition, several components of the intracellular signal-transduction pathways are also regulated by dimerization. For instance, certain cytoplasmic signal-transduction molecules dimerize after activation, and the active form of a transcription factor is often a dimer.[187–190] Thus, protein dimerization has emerged as a general mechanism for the initiation and downstream regulation of signal transduction. The targets listed below that we selected to pursue were chosen not only for their therapeutic importance, but also because each constitutes a distinct stage at which to modulate cellular signaling by controlling such protein–protein or protein–DNA interactions:

- Promote a cell surface receptor homodimerization that is required for activation (EPO agonist).
- Inhibit an extracellular cell surface integrin–protein interaction (MMP-2/$\alpha_v\beta_3$).
- Inhibit an intracellular cell surface integrin–protein interaction (Paxillin/$\alpha_4\beta_1$).
- Inhibit the intracellular heterodimerization of a transcription factor (Myc/Max, LEF-1/β-catenin).
- Inhibit the protein–DNA interaction of a transcription factor (LEF-1/β-catenin, androgen receptor).

The more personal account of our own studies on this series of targets that follows highlights some important aspects of target selection, screening technology, and library generation.

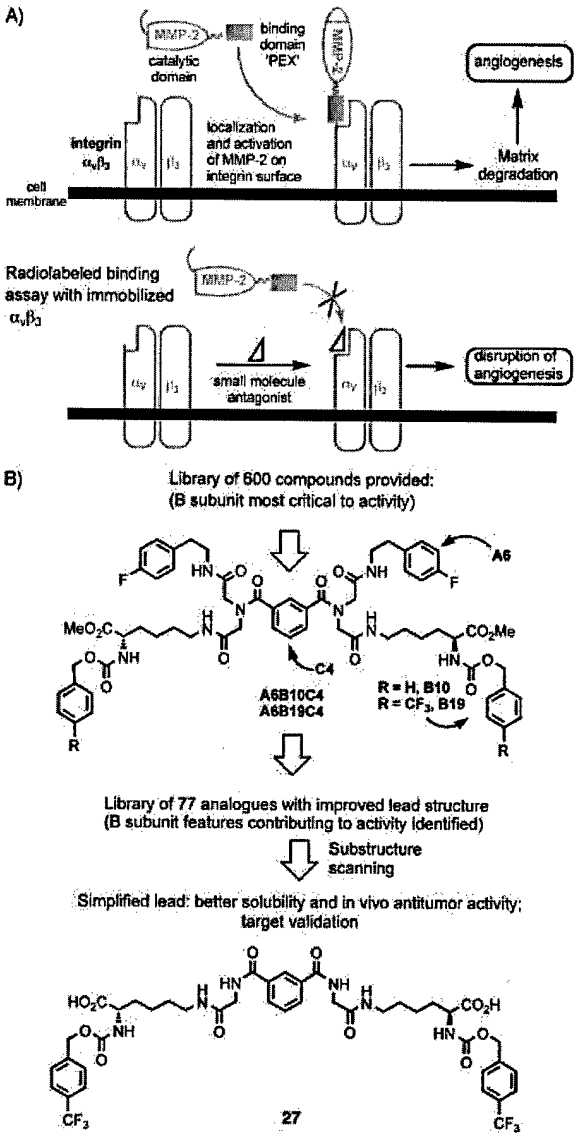


Figure 3. A) Disruption of angiogenesis by blocking matrix metallo proteinase 2 (MMP-2) binding to integrin $\alpha_v\beta_3$. B) Developing the angiogenesis inhibitor **27**.

4.1. MMP-2/$\alpha_v\beta_3$

MMP-2 (gelatinase A), a member of the matrix metalloproteinase family, is secreted by vascular endothelial cells and plays a crucial role in degrading the extracellular matrix during tumor-induced angiogenesis.[226] Consequently, MMP-2 inhibitors have been extensively pursued and developed which diminish tumor growth and metastasis.[226–231] Another key participant in angiogenesis is the integrin[232] $\alpha_v\beta_3$, which mediates cellular interactions with the extracellular matrix.[231,233–239] This heterodimeric cell surface receptor recognizes certain matrix proteins containing the RGD (Arg-Gly-Asp) peptide sequence, which leads to integrin clustering, and ultimately supporting the migration and survival of endothelial cells. Soluble RGD peptides, nonpeptide RGD mimetics, and antibodies that disrupt $\alpha_v\beta_3$ ligation have been shown to inhibit angiogenesis and tumor growth.[240–252] Conservative estimates indicate that at least 40 pharmaceutical firms are developing angiogenesis inhibitors directed at these two targets with at least 27 drugs currently in clinical trials.[231,253–258]

Recently, a novel interaction between MMP-2 and integrin $\alpha_v\beta_3$ was observed by Cheresh et al. when these

proteins were found to colocalize on the surface of angiogenic blood vessels in vivo.[259] Integrin $\alpha_v\beta_3$ binds MMP-2, and this interaction is a requisite step in the cellular utilization of the enzyme on the surface of invasive endothelial cells.[260] A noncatalytic, 193-residue C-terminal fragment of MMP-2, termed PEX, was found to inhibit MMP-2 binding to $\alpha_v\beta_3$ and to indirectly block its cell surface proteolytic activity.[261] PEX was further shown to disrupt angiogenesis and tumor growth in the chick chorioallantoic membrane (CAM). A naturally occurring form of PEX can be detected in vivo in association with cells expressing $\alpha_v\beta_3$, which implicates the role of PEX as an endogenous regulator of angiogenesis. These observations suggested that disrupting the extracellular binding of MMP-2 to integrin $\alpha_v\beta_3$ may be a promising approach to controlling angiogenesis, and could provide a novel therapeutic target for angiogenesis-dependent diseases including cancer (Figure 3).

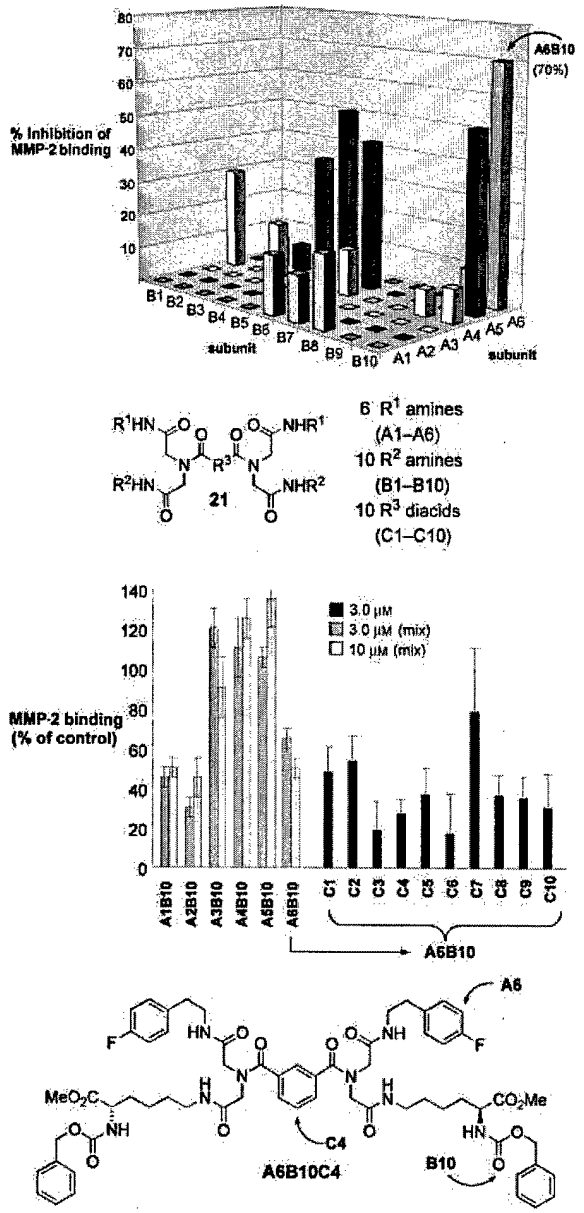


Figure 4. Screening of a 600-compound library that provided the initial lead **A6B10C4** for inhibition of MMP-2 binding to $\alpha_v\beta_3$.

The structural details of the MMP-2/$\alpha_v\beta_3$ interaction are unknown, although it has been established that the integrin binding site is distinct from that which recognizes the RGD sequence in traditional high-affinity ligands (e.g., vitronectin).[260] In the absence of further knowledge of the target structure, a combinatorial chemistry approach was used to identify small-molecule MMP-2/$\alpha_v\beta_3$ antagonists. This approach provided the first members of a new class of anti-angiogenic compounds that derive their activity by inhibiting this protein–protein interaction, and validated this extracellular target for therapeutic intervention.[77,78]

The initial lead came from a library that was screened as 60 mixtures of 10 compounds (Figure 4). The most effective inhibitors contained the **B10** subunit and each of the individual compounds in several of these mixtures was prepared for identification of the active constitutents. Because the 10-compound mixture was generated in the last step of the library synthesis, this entailed simply repeating the last step of the synthesis with individual (instead of mixture) C subunits on the archived samples of the immediate precursor.

Thus, the screening of the compound libraries in a radiolabel binding assay using the immobilized integrin led to the initial identification of **A6B10C4**, its refinement to **A6B19C4** and its subsequent simplification to provide **27** which also possesses improved solubility and in vivo activity (Figure 3). The inhibitor **27** was shown to bind $\alpha_v\beta_3$ and disrupt the binding of MMP-2, but not the integrin ligand vitronectin. It inhibits collagen-matrix degradation without directly inhibiting MMP-2 enzymatic activity, was shown to be a potent inhibitor of angiogenesis in a chick CAM model, and led to near complete reduction in the growth of CS-1 melanoma tumors in vivo in a chick CAM model. Notably, a close analogue of **27** (the benzamide rather than the carbamate) was inactive and used as a control.[77,78]

4.2. Paxillin/α_4: A Successful Use of Positional Scanning

The integrin $\alpha_4\beta_1$ (also know as VLA4, Very Late Antigen 4) is a cell surface receptor that plays an important role in embryogenesis, hematopoiesis, and the immune response.[262–267] It binds natural ligands including vascular cell adhesion molecule 1 (VCAM-1) and an alternatively spliced connecting segment (CS-1) from the extracellular matrix protein fibronectin. It mediates cellular adhesion and activation through a variety of cell–cell and cell–matrix interactions that regulates leukocyte migration into tissues during inflammatory responses and lymphocyte trafficking.[262–267] This integrin is believed to regulate cellular functions differently from other integrins because the α_4 cytoplasmic tail binds tightly to the signaling adapter protein Paxillin through a short conserved sequence motif dominated by two residues (Glu 983 and Tyr 991).[268] The α_4/Paxillin

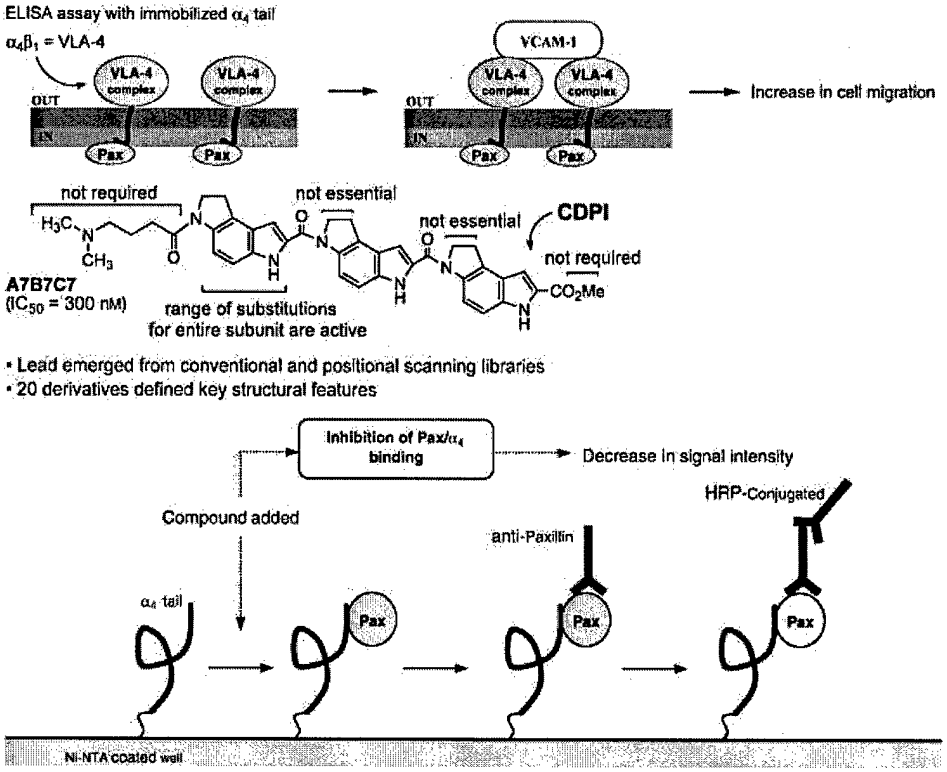


Figure 5. Top: Inhibition of cell migration by blocking intracellular paxillin/α4 binding, middle: identification of the lead structure **A7B7C7** through high-throughput screening using an ELISA assay and bottom: description of the high-throughput screen (ELISA). VLA-4 = very late antigen 4, Pax = paxillin, HRP = horseradish peroxidase

interaction leads to enhanced rates of cell migration and reduced rates of cell spreading, focal adhesion, and stress fiber formation (Figure 5).[269,270] These biological responses to integrin-mediated cell adhesion contribute to leukocyte migration and changes in gene expression important in chronic inflammation. While extracellular inhibitors for this class of integrins are under development for the treatment of asthma and multiple sclerosis,[271–273] inhibitors targeting the unique cytoplasmic Paxillin/α_4 interaction had not been explored.

The target protein–protein interaction again represents one in which no structural information is available and no existing inhibitors had been disclosed. The examination of our libraries (ca. 40000 compounds) in an ELISA assay enlisting the immobilized α_4 cytoplasmic tail and screening for inhibition of Paxillin binding provided the first and potent inhibitors of this protein–protein interaction (Figure 5). The preparation of key substructure analogues of **A7B7C7** identified structural features required for activity, those available for further modifications, and those which may be removed or modified without significantly changing the activity. The most potent lead structure was shown to inhibit $\alpha_4\beta_1$-mediated human Jurkat T cell migration in a dose-dependent manner validating the intracellular Paxillin/α_4 interaction as a useful and unique target for therapeutic intervention. Notably, such target validation is often first established by biological techniques or with monoclonal antibodies or peptide consensus sequences derived from the receptor or its ligands. Consequently, it is of note that even the initial target validation as well as its suitability for small-molecule therapeutic intervention came from the screening of small-molecule libraries. Moreover, the lead structure emerged from a library of 1000 compounds that had been prepared in two formats: 1) a traditional small-mixture format composed of 100 mixtures of 10 compounds, and 2) a less traditional positional-scanning library. Parallel testing of both libraries provided the rare opportunity to critically compare the results derived from the two approaches.[79]

Thus, an important feature in this work rested with the parallel screening of a positional-scanning library for which every component contained in a traditional compound library was also present, but assembled such that immediate deconvolution was possible. Scanning for the best subunit at position A by screening sublibrary **AxBC** showed that the greatest inhibition was observed for mixture **A7BC**, identifying the CDPI subunit as the best subunit at position A (Figure 6). Using the same procedure, CDPI was also identified as the most effective subunit at positions B and C (Figure 6). Immediate deconvolution of the results identified the **A7B7C7** combination as a potent Paxillin/α4 binding antagonist. This same lead compound (**A7B7C7**) was identified from the simultaneous and more traditional library screening. This unbiased screening success must be tempered by the fact that the positional-scanning-library screening and

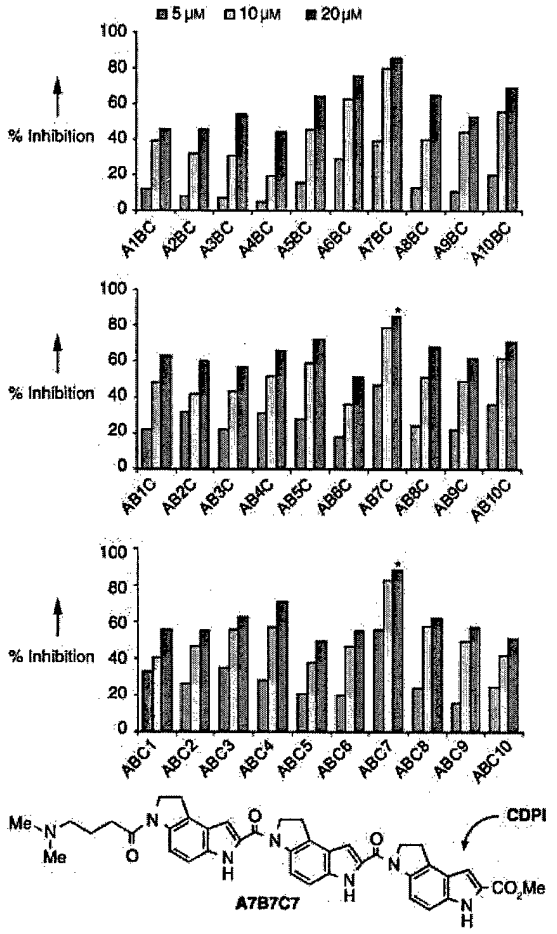
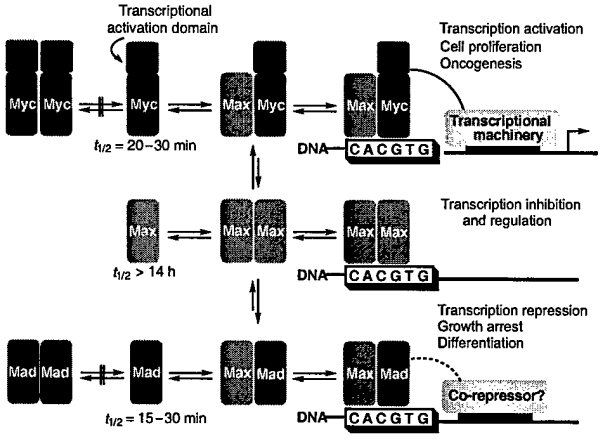


Figure 6. Inhibition of paxillin/α4 binding by the positional scanning libraries AxBC (top, A scan), AByC (middle, B scan), and ABCz (bottom, C scan). Each mixture of 100 compounds was tested in triplicate at 5, 10, and 20 μM (total compound).

direct deconvolution did not identify all candidate inhibitors which were identified from the traditional library. This situation is a natural consequence of testing larger 100-compound mixtures and the relative insensitivity of the assay to the contribution of any single, uniquely acting compound in the mixture. Thus, the global observations were effectively detected with the positional-scanning library, and a useful lead structure with defined properties was identified.[79] However, more subtle discoveries within the libraries were not identified. Thus, the disadvantages associated with the loss of their detection and this information contained within the library must be balanced against the advantages of the ease of synthesis of the parent libraries and judged in light of the objectives of the library screening. As in this case, the positional-scanning libraries typically would be only effective for lead identification and would be less suitable for lead optimization.

4.3. Myc/Max

The *c-myc* proto-oncogene is involved in the progression of a wide-range of human tumors.[274–284] Proteins in the immediate Myc network are essential regulators of cell growth and differentiation. Activation and transformation occur mainly through elevated expression of its gene product Myc, a short-lived nuclear protein and an important member of the basic helix-loop-helix leucine zipper (bHLHLZ) transcription factors (Figure 7).[285,286] The oncogenic potential of *c-myc* has been demonstrated in transgenic animals, reconstituted organs or cell culture, and typically requires the action of at least one additional oncogene, such as *ras* or *bcl-2* (inhibits apoptosis).[274–284] In normal cells, Myc is required for cell proliferation and prevents differentiation. It is transiently expressed in response to mitogenic stimuli and it has a short half-life ($t_{1/2}$ = 20–30 min). Its aberrant expression in the



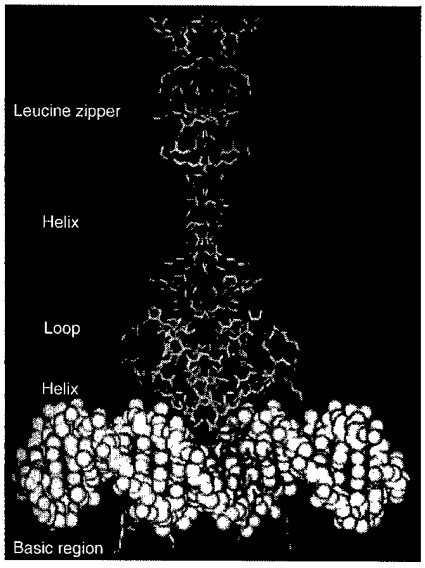


Figure 7. Top: Myc-Max-Mad network of transcription control: cell proliferation versus differentiation Max = Myc associated factor X. Bottom: Max(22-113) dimer complexed with DNA. Reproduced with permission from *Nature*, reference [285].

absence of such growth factors drives cells into cell cycle but also induces apoptosis unless co-oncogenic mechanisms are present (i.e., inhibition by constitutive *bcl-2* overexpression).

All known activities of Myc require heterodimerization with Max,[285–293] a constitutively expressed bHLHLZ protein that is stable ($t_{1/2} = > 14$ h) and always present. The Myc/Max heterodimer binds the DNA sequence CACGTG (E box element) and activates transcription through the transactivation domain of Myc. DNA binding occurs through the basic region while both the HLH and the leucine zipper form the dimerization interface.[285,286] Max, but not Myc, forms homodimers which bind the same DNA site and inhibit transcription. In part, this results from the competition for Max and for the common DNA site. In addition, Max heterodimerizes with another family of bHLHLZ proteins that include Mad (Mad-1), Mxi-1, Mad-3, and Mad-4 that are transiently expressed ($t_{1/2} = 15$–30 min) and signal growth arrest and differentiation.[296,297] Their dimerization behavior is similar to that of Myc in that they do not homodimerize or interact with Myc family members, but readily form heterodimers with Max that bind the CACGTG core consensus sequence. Mad and Myc compete for Max with approximately equal affinities and the heterodimers are formed in preference to Max homodimerization. Mad/Max heterodimers repress or antagonize Myc/Max transcriptional activation. In normal cells, the expression of Mad is induced by differentiating stimuli while Myc expression is repressed.[274–283] Thus, Myc/Max dimers induce proliferation and apoptosis in nontransformed cells while Max/Max and Mad/Max are involved in growth arrest, differentiation, and cell survival. At the center of this network lies Max which is metabolically stable ($t_{1/2} = > 14$ h) and constitutively expressed. Its heterodimer partners, Myc ($t_{1/2} = 20$–30 min) and Mad ($t_{1/2} = 15$–30 min), are rapidly degraded and highly regulated. Mitogenic stimulation induces a rapid rise in Myc levels, and a shift in the equilibrium to Myc/Max dimers and transcriptional activation. In transformed cells, when accompanied by oncogenic events inhibiting apoptosis, constitutive Myc expression maintains the cell proliferation state and prevents growth arrest and differentiation.[274–283] Thus, the function of the Myc-Max-Mad network relies on four major features:

1) Max is a required dimerization partner for Myc (or Mad) to bind DNA and execute its biological functions.
2) Myc/Max and Mad/Max heterodimers form in preference to Max/Max homodimers.
3) The highly regulated expression levels of Myc and Mad induce rapid changes in the equilibrium between the dimers. The Myc/Max heterodimer activates transcription throughout all stages of the cell cycle and maintains the cell in a state of cell growth and proliferation while the Mad/Max heterodimer signals an exit from the cell cycle, growth arrest, and differentiation.
4) All dimers compete for common target DNA sites, and the Myc/Max transcription activation is antagonized by the other dimers. Thus, the aberrant expression of Myc that leads to transformation may be addressed through inhibition of Myc/Max heterodimerization and its subsequent transcriptional activation. This provides an attractive therapeutic target for the treatment of cancer that is central to the disease.

A novel FRET assay for monitoring Myc/Max dimerization was developed specifically for this work (Figure 8). Thus, the fusion proteins (MycCFP and MaxYFP) consisting of the

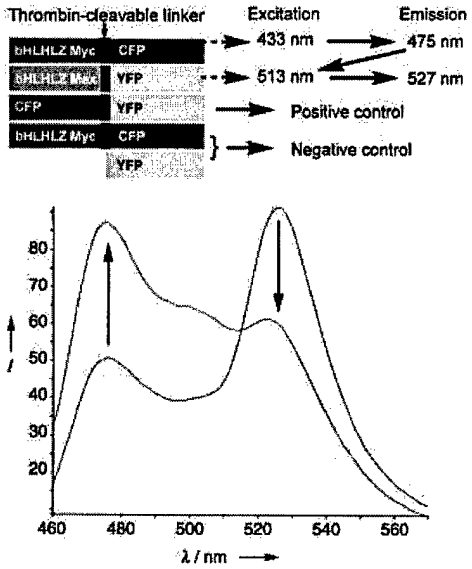


Figure 8. Fluorescence resonance energy transfer (FRET) assay. Addition of dimerization inhibitors induces a decrease in the intensity at 527 nm (dimer) and an increase in intensity at 475 nm (monomer). Compounds that give rise to a large ratio I_{475}/I_{527} are selected. CFP = cyan fluorescent protein, YFP = yellow fluorescent protein.

bHLHLZip domain of human *c-myc* linked to the N-terminus of cyan fluorescent protein (CFP) and the bHLHLZip domain of Max fused to the N-terminus of yellow fluorescent protein (YFP) were cloned, expressed, and purified for use in the assay. Inhibition of dimerization is measured as a loss of the FRET signal emission at 527 nm (YFP) following excitation of CFP at 433 nm. Using this assay, our library of 40000 compounds was screened for inhibition of Myc/Max dimerization. Two closely related leads emerged from an isoindoline library of 240 compounds (Figure 9). These remarkably small inhibitors (38% FRET inhibition at 25 μM for **28**) were subsequently shown to inhibit Myc/Max binding in a conventional ELISA assay and to inhibit Myc/Max DNA binding in a gel-shift assay confirming the observations made through use of the novel FRET assay. Impressively, the two most potent leads were shown to inhibit Myc-induced transformation of chicken embryo fibroblasts ($IC_{90} = 20$ μM) establishing functional activity of the compounds in a cellular assay.[80]

4.4. LEF-1/β-catenin

The majority of colorectal tumors contain mutations in the tumor-suppressor protein, adenomatous polyposis coli

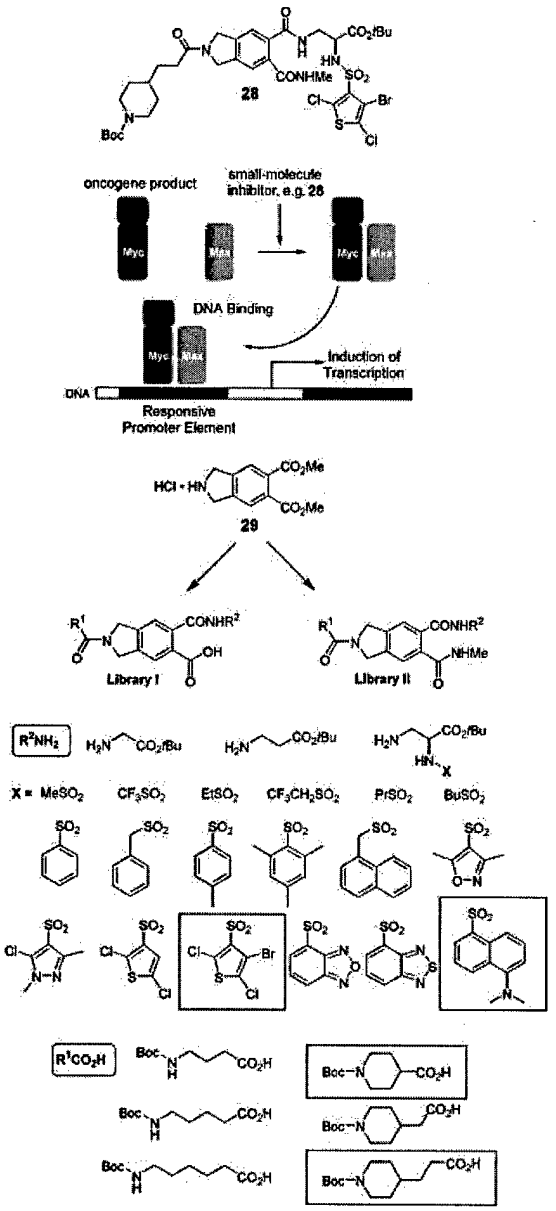


Figure 9. Inhibition of Myc-Max heterodimerization and aberrant gene transcription: antitumor target validation.

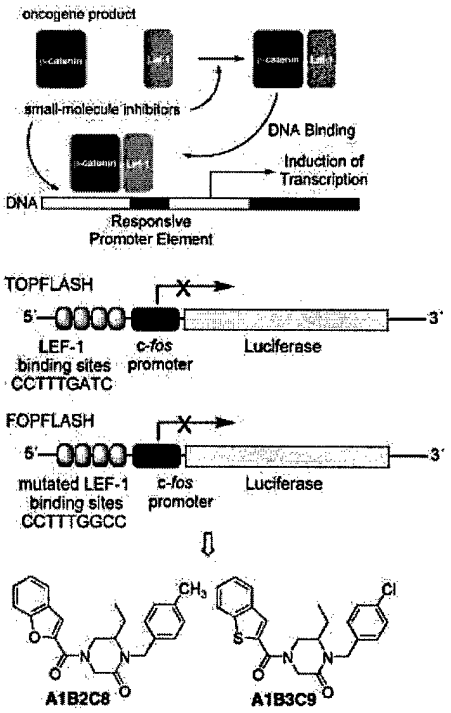


Figure 10. Inhibition of LEF-1/β-catenin-mediated gene transcription: controlling aberrant gene transcription and antitumor target validation. More than 40000 compounds were screened for inhibition of LEF-1 mediated gene transcription in a TOPFLASH luciferase reported assay. Two leads **A1B2C8** and **A1B3C9** exhibit TOPFLASH selectivity, they are derived from a 150-member non-amide based IDA library of piperazinones.

(APC) of the colon, that result in the release and accumulation of β-catenin in the cell nucleus. β-Catenin binds to and activates transcription factors including LEF-1 (Lymphoid-enhancers binding factor 1).[298-301] This results in upregulated and aberrant gene expression which is a key step in the development of colon cancer (Figure 10). The LEF-1 transcription factors (TCF) that interact with β-catenin share an identical DNA-binding domain referred to as the high mobility group (HMG) domain recognizing the sequence 5′-CTTTGWW-3′ (W = A or T). Importantly, LEF-1 binds the DNA minor groove[302] through its HMG domain making it an ideal target for libraries of DNA minor-groove binding ligands. When β-catenin interacts with the amino terminus of LEF-1, it activates transcription. In the absence of β-catenin, LEF-1 actively represses transcription. It is thought that LEF-1 bound β-catenin displaces corepressors and β-catenin binds LEF-1 through a domain distinct from that of the corepressors. The oncogenic potential of β-catenin through its interaction with LEF-1 has been demonstrated by Vogt et al.[303]

Initial studies were conducted screening our library of 40000 compounds enlisting a TOPFLASH/FOPFLASH cell-based reporter assay[298] transiently transfected into a colon cancer cell line (SW480) that has an APC mutation which allows β-catenin to accumulate in the nucleus. The use of a luciferase reporter assay permits the detection of leads that act by disrupting the LEF-1/β-catenin binding interaction or that inhibit LEF-1/DNA binding. This approach led to the identification of two classes of small molecules that inhibit LEF-1/β-catenin-mediated gene transcription. The first class, of which **A1B2C8** (3.7-fold inhibition at 10 μM) and **A1B3C9** (4.8-fold inhibition at 10 μM) are representative, proved to be a surprising class of small molecules that probably do not act by inhibiting LEF-1/DNA binding, but rather by inhibiting LEF-1/β-catenin binding (Figure 10). Their evaluation in a functional cellular assay revealed that they inhibit the β-

catenin-induced tumor cell transformation of the colon cell line SW480. The initial library of 150 members from which the lead emerged was expanded to more than 350 members with the identification of structural features important for activity in the reporter and functional assays.[71] These studies are now being expanded to larger libraries based on **A1B2C8** and **A1B3C9** and with the development of a FRET assay capable of directly screening for disruption of LEF-1/β-catenin binding. This assay will serve to define the site of action of the leads related to **A1B2C8** and should provide new candidate lead structures through the screening of future libraries. The second class of selective TOPFLASH inhibitors emerged from a library of minor-groove DNA binding compounds and are discussed in detail in Section 5.3.

4.5. EPO/EPOR: Agonists from Antagonists

Erythropoietin (EPO), a 34-kDa monomeric glycoprotein composed of 165 amino acid residues and an almost equivalent mass of carbohydrate, is the principle factor regulating red blood cell production.[304–311] EPO exerts its effects by controlling the proliferation, differentiation, and maturation of erythroid progenitors. The EPO gene, which shows no homology with other known genes, is a single-copy gene and highly conserved across different species. Recombinant human EPO has found clinical applications in the treatment of anemias particularly those resulting from renal disease including kidney damage, anemia resulting from malignancy, AIDS or AIDS treatment, and anemia associated with myelodysplastic syndromes.[308–311] While EPO is presently the most important biotechnology product and its use has evolved into a $>$\$7 billion/year market, its administration requires intravenous or subcutaneous injection. This situation has led to an interest in developing small-molecule mimics of EPO to replace the recombinant human protein. More recently, EPO itself is thought to be a potentially useful antitumor agent,[312] although the mechanism is not yet well understood. Prompted by observations that EPO-treated patients experience longer survival times and more complete remissions than untreated patients, the in vivo evaluation of EPO itself for antitumor efficacy was successfully conducted.[312] EPO elicits its response through binding and homodimerization of its receptor EPOR[313–316] which is a member of the Class I cytokine receptors (Table 4).[183–185] The Class I cytokine receptors are characterized by the presence of one or two conserved 200-amino acid fibronectin type III-like domains, four conserved cysteine residues, and the conserved motif Trp-Ser-Xaa-Trp-Ser (WSXWS) in the extracellular domain. The intracellular domains of the cytokine receptors, unlike the receptor tyrosine kinases, lack intrinsic enzymatic activities. Ligand binding induces dimerization or oligomerization and this allows the interaction (dimerization) and activation (cross phosphorylation) of noncovalently associated tyrosine kinases (Janus Kinases; JAKs), Figure 11.[47,71,188,318] In turn, the activated JAKs phosphorylate cytoplasmic domain tyrosines of the receptor enlisted to recruit signaling proteins including STATs (Signal Transducers and Activators of Transcription).[187,188] Recruitment of STATs results in their phosphorylation, homo- or heterodimerization, and subsequent translocation to the nucleus where they directly activate gene transcription. This is the most direct signal-transduction pathway known and beautifully enlists coordinated protein–protein homo- or heterodimerization of signaling molecules to link the initiation receptor-dimerization event directly to gene transcription. Notably, receptor homodimerization enlists JAK and STAT homodimers (JAK2/JAK2 and STAT5/STAT5 for EPOR) that bind a unique DNA sequence and upregulate the target genes, whereas receptor heterodimerization enlists JAK and STAT heterodimers (i.e., STAT4/STAT5) that target a different DNA sequence for upregulating a different set of targets genes. Thus, this combinatorial dimerization strategy is enlisted in the natural signaling cascade to maximize the diversity of signals with a limited number of proteins.

The approach we adopted is generalizable to all cytokine receptors. Because of the scale on which the synthesis is conducted, the libraries could be enlisted to screen for agonists or antagonists against multiple receptors. Although not an objective of present work, their availability should allow such studies in the future. As detailed earlier, homo- or heterodimerization linkage of ligand binding antagonists can be enlisted to develop ligand agonists.[170] Potential applications of such studies with the cytokine receptors are summarized in Table 4.

Table 4: Class I cytokine receptors and approved or potential therapeutic applications of cytokine agonists and antagonists.

Class I cytokine receptors		
Family	Examples	Activation Characteristics
GH receptor	GHR, EPOR, TPOR, PRLR, G-CSFR	homodimers
IL-3 receptor	IL-3R, GM-CSFR, IL-5R	heterodimerization with β_c
IL-6 receptor	IL-6R, LIFR, CNTFR, IL-11R	heterodimerization with gp130
IL-2 receptor	IL-2Rα, IL-2Rβ, IL-4R, IL-7R	heterodimerization with IL-2Rγ
Therapeutic applications of cytokine agonists and antagonists		
Cytokine	Agonist	Antagonist
EPO	anemias, selective blood donation, cancer	
TPO	thrombocytopenia	
IL-2	cancer	histoincompatibility
IL-3	leukopenia, myeloid reconstitution	leukemia
IL-4	inflammation, cancer	allergy
IL-6	thrombocytopenia	cancer, osteoporosis, inflammation
IL-11	thrombocytopenia	
IL-12	cancer, infections	histoincompatibility, autoimmunity
G-CSF	neutropenia, myeloid reconstitution	leukemia
GM-CSF	leukopenia, myeloid reconstitution	leukemia
IFN α/β	cancer, viral infections, autoimmunity	inflammation
IFN γ	chronic granulonatous disease, infections	inflammation, autoimmunity

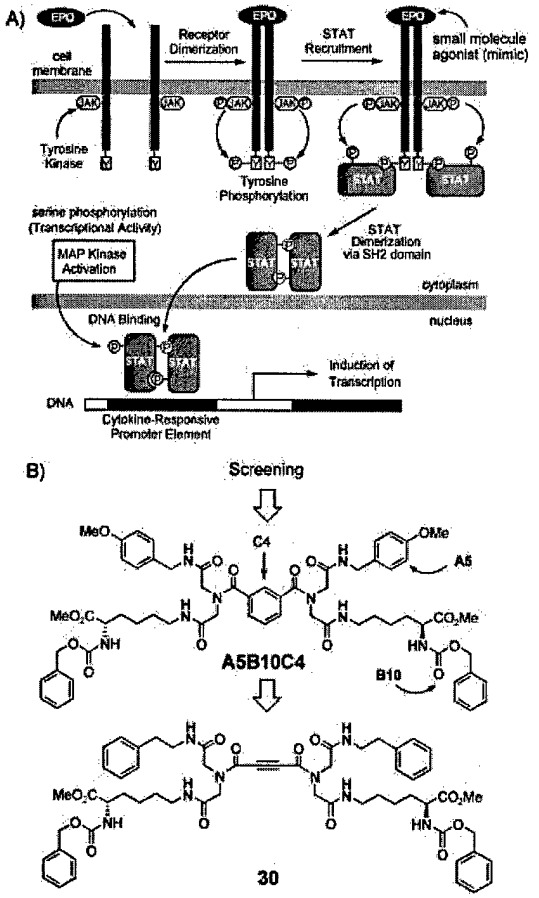
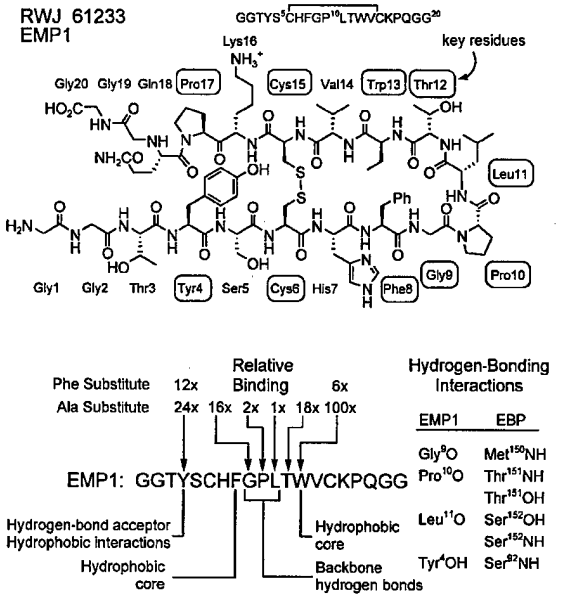


Figure 11. Top: EPO agonists (mimics) that function by promoting EPOR dimerization. Bottom: Development of EPO agonist **30** which is active in cellular functional assay, step 1: over 100000 molecules were screened for EPOR binding, the B subunit is shown to be the most critical for binding activity. Step 2: A library of 115 A analogues provided an improved lead structure. The B subunit features contributing to activity were identified (77 analogues). A library of 70 C analogues provided an improved lead structure, three 1000-member libraries with alternative templates were examined. Higher order IDAs examined: binders examined in functional assay.



Conserved motif in EPO mimetic peptides: xxxYxCxGPxTWxCxPxxx
EMP1-EBP interaction: 67% hydrophobic; 33% polar

Figure 12. Structure of EPO mimetic peptide (EMP1), summary of relative EPOR binding, and summary of key contacts derived from the X-ray crystal structure of the complex with EBP (EPOR).

The disclosure of a 20-amino acid cyclic peptide RWJ 61233 (EMP1) identified in a phage library that binds to and activates EPOR ($K_D = 200$ nM versus $K_D = 200$ pM for EPO)[319,320] along with the X-ray structure determination of the peptide–receptor complex[321] revealed in detail how two molecules of the peptide dimerize EBP (EPO binding protein, the extracellular domain of EPOR; Figure 12 and 13).[322] The structure of the complex revealed a peptide dimer flanked by two molecules of the EBP and also reveals a striking structural homology to the growth hormone and prolactin receptors.[323,324] Interestingly, in the stem regions the two molecules of EBP have very little contact (ca. 75 Å^2) and, contrary to expectations, the WSXWS motif is not involved in ligand binding nor located in or near the ligand binding site or dimerization interface. Consistent with mutational studies of growth hormone and its receptor[325–328] which revealed that the majority of the intermolecular binding stabilization is derived from a few residues in the contact interface, the dimeric peptide binds EBP drawing upon a limited number of contacts for binding and dimerization. Further support for the key contacts has been derived through modifications of the cyclic peptide (Figure 12).[319,320] This result suggested that relatively small ligands may be devised as EPO agonists even though the cyclic peptide dimer itself is quite large. The cyclic peptide contains a single disulfide bond Cys 6/Cys 15 which links two short β-strands (residues 4–7 and 13–16) that are connected by a type I β-turn (Gly 9-Pro 10-Leu 11-Thr 12).[329] The dimerized cyclic peptide forms a hydrophobic clamp that slides into a receptor dimerization interface making contacts with both receptor molecules. Central to this binding is a hydrophobic interface, Pro 149-His 153, which houses residues engaged in main-chain and side-chain hydrogen bonds to Gly 9-Thr 12/Trp 13 of the peptide. Critical and also central to the binding dimerization is the receptor Phe 93 which provides key contacts with Trp 13 and Tyr 4 complementing the adjacent Ser 92/Tyr 4-OH hydrogen bond. Similarly, Phe 8 spans the binding sites defined by Leu 33, Met 150, and Phe 205 which contributes significantly to the stabilization of the dimer complex. Our candidate library agonists were selected on the basis of maintaining and spanning these critical interactions. Although space precludes a detailed description of the modeling efforts that went into the design of our initial libraries, it is perhaps sufficient to indicate that effective EPOR binding ligands were obtained in the first libraries and their further exploration has provided even better binders.

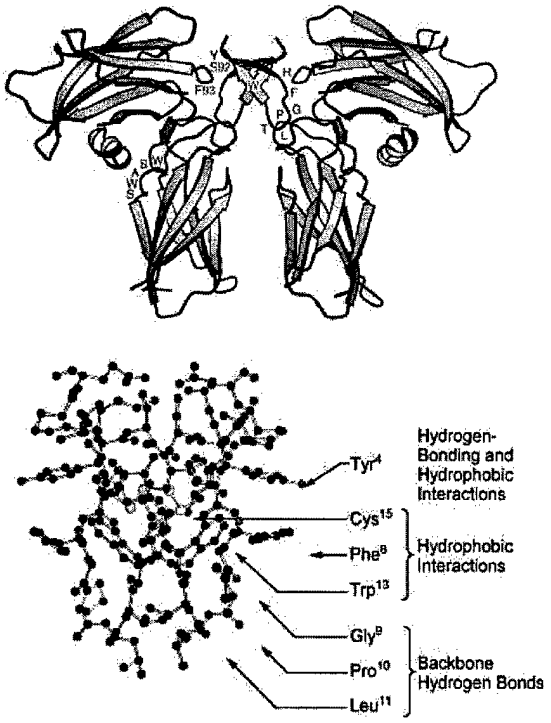


Figure 13. Top: X-ray structure of the EBP-EMP1 complex. Bottom: Key residues involved in the peptide–EBP interaction.

Employing libraries prepared from *N*-Boc-iminodiacetic acid anhydride and a three- or four-step solution-phase reaction with acid–base extractions for purification, numerous libraries were prepared (roughly 100 000 compounds) and evaluated for EPOR binding. A number of EPOR binding agents were identified that inhibited I^{125}-EPO binding to EPOR (IC_{50} = 1–80 μM). These antagonists of EPO binding were covalently linked to provide C_2 symmetric dimers capable of receptor dimerization and activation. Several of the most active compounds were found to be weak agonists and induced concentration-dependent proliferation of an EPO-dependent cell line (UT-7/EPO) exhibiting the characteristic bell-shaped dose response curve while having no effect on a similar cell line (FDC-P1) lacking the EPOR (Figure 11 and 14).[82]

To date, only three reports of EPO agonists have been disclosed[82,319,330] and those reported above[82] are only about 10-fold less potent than EMP1[319] (IC_{50} = 5–10 versus 0.5 μM). They are substantially more efficacious than a recently reported Merck agonist[330] (15–20% versus 4% of activity of EPO), considerably smaller than EMP1 ($M_r \approx 2100$, but functions as dimer of $M_r \approx 4200$) or the Merck agonist ($M_r \approx$ 6400), and are substantially more accessible (four synthetic steps). Key binding elements have been identified allowing for further potential reductions in the size of the identified leads.

Since the agents act on the extracellular domain of the receptor, they need not be permeable to the cell. As such, the relative size of the agonist structures need not be as significant a concern as it is with an intracellular target. Similarly, one can anticipate that the amide bonds present in the existing leads can be engineered out of the structures (Figure 14). Finally, one can anticipate that the size of the agonist structures may be reduced and their affinity increased through examination of the X-ray structures of complexes with EBP. These studies should allow the design of additional key interactions into the leads, the trimming away of non-productive elements of the compounds, and permit the design of rigid structures embodying the EBP bound conformation. Finally, it is now clear that simple dimerization of the EPO receptor is insufficient for receptor activation.[170,316,331–334] This dimerization, or conformational reorganization of a cell surface receptor dimer, must occur in a manner that

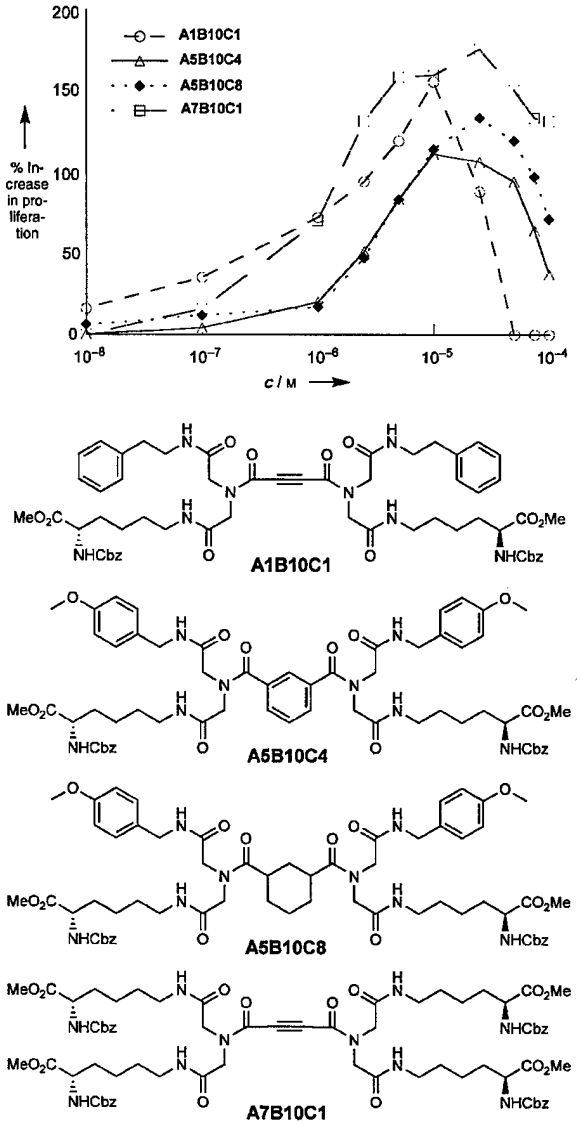
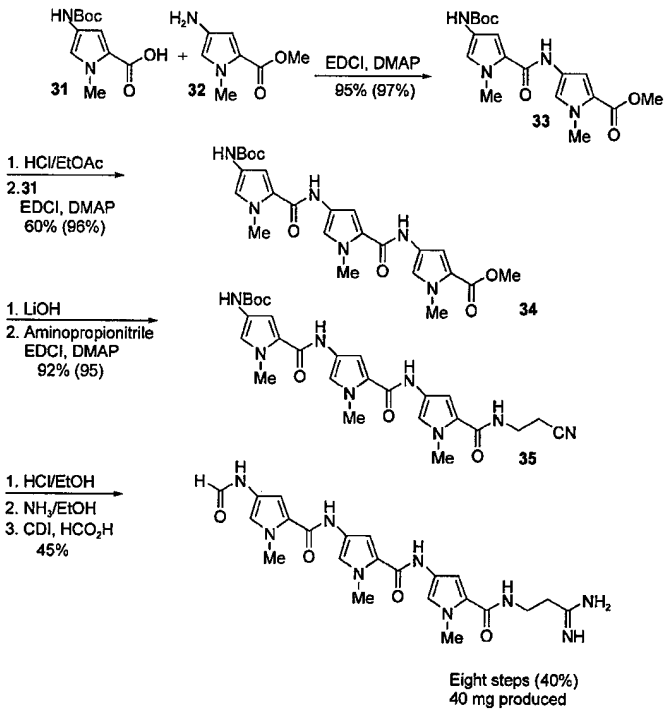


Figure 14. Concentration-dependent proliferation activity in an EPO-dependent cell line exhibiting a bell-shaped dose-response curve.

brings the receptor cytoplasmic domains in close proximity to one another permitting the JAK cross phosphorylation and initiation of the signaling cascade. Thus, rigid variants of the lead structures that not only bind the ligand site and stabilize receptor dimerization, but that also more effectively induce the active conformation of the receptor dimer may help convert the partial agonists into full agonists.

5. Targeting Protein–DNA Interactions

Studies on the high-throughput synthesis and screening of DNA binding compounds[72,84–88] have also been detailed that are related to interests in understanding and exploiting their properties.[89–96] Our approach integrates the solution-phase techniques for the synthesis of libraries with a technique we introduced for rapid high-throughput screening for DNA binding affinity or sequence selectivity.[87,97,98] These techniques can be combined to rapidly explore and define the structural features responsible for the sequence-selective DNA binding properties of known agents, to discover new patterns for small-molecule recognition of DNA (new base-pair (bp) codes), and to screen for compounds that selectively target consensus sequences of transcription factors for controlling aberrant gene transcription (i.e., LEF-1/β-catenin, androgen receptor).



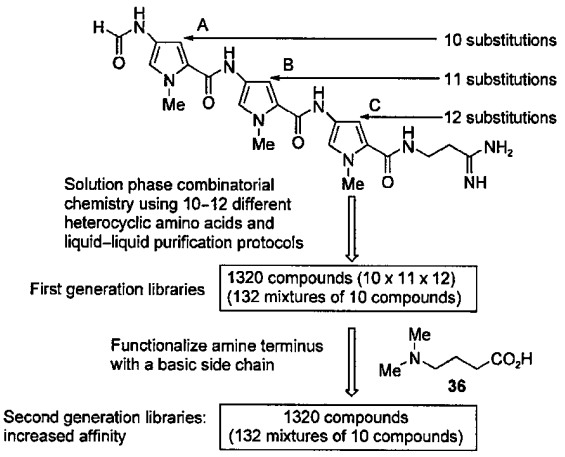
Scheme 8. Solution-phase synthesis of distamycin A. Optimized yields in parentheses. All the compounds were obtained in >95% purity after acid–base extraction, Chromatographic purification was not used. DMAP = 4-*N*,*N*-dimethylaminopyridine, Boc-Pyr-OH = **31**, EDCI = 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide, CDI = 1,1′-carbonyldiimidazole.

The regulation of gene expression is based on the sequence-selective recognition of nucleic acids by repressor, activator, and enhancer proteins. Selective control of such processes has been a long-standing goal and small molecules that selectively bind DNA and activate (block a repressor) or inhibit (block an activator) gene expression hold great promise as therapeutics.[335–356] Complicating such studies is the recognition that it is often not a single sequence that is ideally targeted, but rather an ensemble of related sites that compose the consensus binding sequence of a nuclear receptor or transcription factor. Consequently, the discovery of such agents has been slow owing to the complexity associated with understanding small-molecule/DNA interactions, the iterative process of designing and synthesizing individual compounds targeted towards specific DNA sequences, and the technically demanding techniques involved in the determination of their binding affinity for any given sequence, much less an ensemble of related sequences.

5.1. High-Throughput Synthesis of DNA-Binding Compounds Using Solution-Phase Library Techniques

5.1.1. Distamycin and a Library of 2640 Analogues

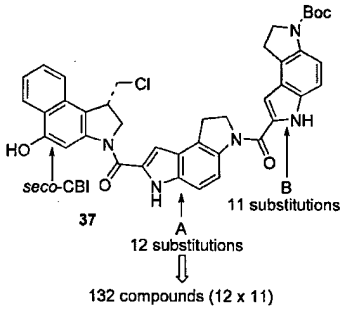
A rapid, parallel synthesis of distamycin A analogues was developed enlisting the simple, acid–base liquid–liquid extraction for purification and isolation of each intermediate and final product (Scheme 8). The utility of this method was first demonstrated with the preparation of distamycin A (eight steps, 40% overall yield), a naturally occurring antitumor antibiotic, where each of the reaction products was isolated in at least 95% purity.[87] A prototypical library of 2640 analogues was assembled in a small-mixture format of two libraries of 132 mixtures of 10 compounds providing each compound in multimilligram quantities sufficient for repetitive screening in multiple assays (Scheme 9). Screening the initial library in a functional assay for cytotoxic activity (L1210) revealed two uniquely active compounds which were found to be 1000-times more potent than distamycin A.[87] A complementary rapid, high-throughput screen for DNA binding affinity was also developed which is applicable for assessing binding to DNA homopolymers or specific sequences (discussed in Section 5.2). Using this technique, alternative AT-rich binding agents were identified establishing the usefulness of the approach and providing two new and effective DNA-binding distamycin A analogues. In addition, a comparison of several distamycin analogues established substituent contributions to AT-rich binding that may be safely implemented in future libraries. Extension of these studies to identify effective binders to predefined sequences was conducted in the context of the androgen response elements PSA-ARE-3 and the ARE-consensus sequences, the latter of which contains an AT-rich sequence interrupted by a GC base pair, and for which effective binders might prove useful in the treatment of resistant prostate cancer.



Scheme 9. Solution-phase synthesis of a library of 2640 potential DNA-binding distamycin analogues.

5.1.2. *CC-1065 and the Duocarmycins: A 132-Membered Library Defining the Role of the DNA-Binding Domain*

A library of 132 CBI (1,2,9,9a-tetrahydrocyclopropa[*c*]-benzo[*e*]indol-4-one) analogues of CC-1065 and the duocarmycins, potent antitumor antibiotics that alkylate duplex DNA, was prepared utilizing the solution-phase technology of acid–base liquid–liquid extraction for the isolation and purification steps (Scheme 10).[88] The 132 analogues consti-



Scheme 10. Solution-phase synthesis of 132 CC-1065 analogues based on CBI. All the compounds were purified by liquid–liquid acid–base extraction.

tuted a systematic study of the DNA binding domain with the incorporation of dimers composed of monocyclic, bicyclic, and tricyclic (hetero)aromatic subunits. From their examination, clear trends in cytotoxic potency and DNA alkylation efficiency emerged highlighting the importance of the first-attached DNA-binding subunit (A subunit): tricyclic is more active than bicyclic, which in turn is far more active than monocyclic (hetero)aromatic subunits (Figure 15). Notably, the trends observed in the cytotoxic potencies paralleled those observed in the relative efficiencies of DNA alkylation.

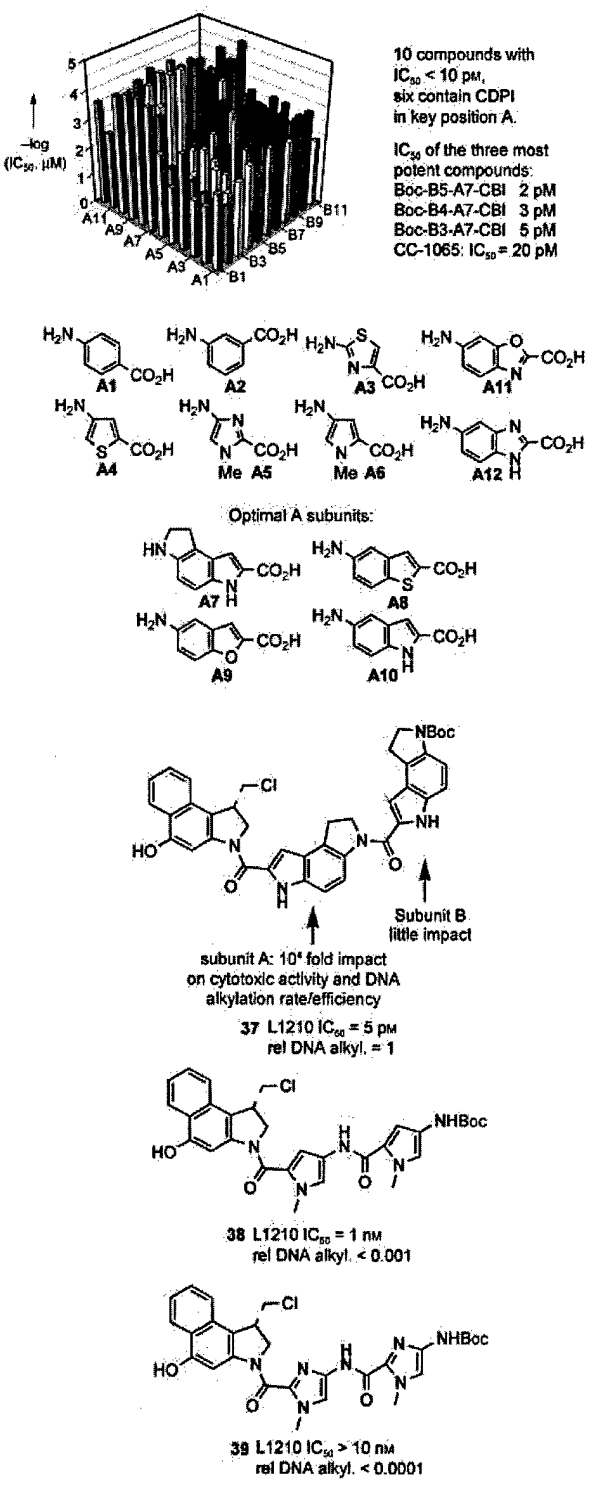


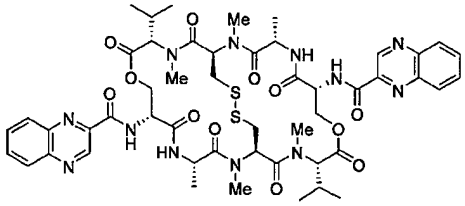
Figure 15. L1210 Cytotoxic activity of the CC-1065 analogues.

Our interpretation of these results is that the trends represent the partitioning of the role of the DNA-binding subunit(s) into two distinct contributions. The first of which is derived

from an increase in DNA-binding selectivity and affinity which leads to property enhancements of 10–100-fold and is embodied in the monocyclic series. The second, which is additionally embodied in the bicyclic and tricyclic heteroaromatic subunits, is a contribution to the catalysis of the DNA alkylation reaction that provides additional enhancements of 100–1000-fold. The total enhancement thus exceeds 25000-fold. Aside from the significance of these observations in the design of future CC-1065/duocarmycin analogues, their importance to the design of hybrid structures containing the CC-1065/duocarmycin alkylation subunit should not be underestimated. Those that lack an attached bicyclic or tricyclic A subunit, that is, duocarmycin/distamycin hybrids, can be expected to be intrinsically poor or slow DNA alkylating agents.

5.1.3. A Prototype Triostin A Analogue Solution-Phase Synthesis

A solution-phase synthesis of azatriostin A, a key analogue of the DNA bisintercalator triostin A (Scheme 11), was developed such that libraries of triostin A analogues could be



Scheme 11. Triostin A: naturally occurring DNA bisintercalator.

prepared.[72] This work, in conjunction with the distamycin A and CC-1065 analogue libraries, highlights the ease with which natural product libraries targeting DNA can be constructed using solution-phase technology to rapidly identify the key structural components of the natural product as well as to discover and develop new therapeutics.

5.2. The FID Assay: A High-Resolution, High-Throughput Assay for Establishing DNA-Binding Affinity and Selectivity

A variety of techniques are commonly used to establish the DNA-binding properties of small molecules.[357] Most are technically challenging and require the knowledge of specialized biochemical procedures and assay reproducibility comes only with experience.[358,359] Prior to our efforts, none were applicable to the high-throughput screening required for assaying a library of compounds against an ensemble or complete library of sequences. The most widely utilized methods are footprinting or affinity cleavage and both have been used qualitatively and quantitatively to establish DNA-binding selectivity or affinity. Intrinsic in the methods employed in the destructive assays is the binding characterization at the highest affinity sites within an incomplete set of sites contained in a 100–200-bp segment of DNA, and modest- or low-affinity sites are not easily probed. As a consequence, the full DNA-binding profile of candidate ligands is not easily assessed or compared using these techniques.

The nondestructive fluorescent intercalator displacement (FID) assay, which addresses these issues of ease of use and information content, utilizes the displacement of ethidium bromide (or thiazole orange) from hairpin deoxyoligonucleotides (Figure 16).[72,87] A hairpin deoxyoligonucleotide con-

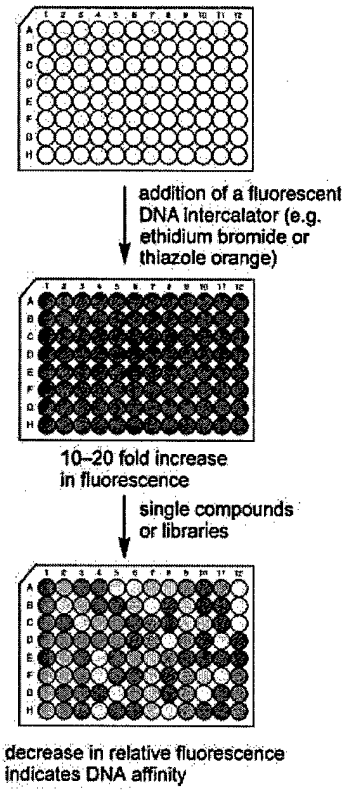


Figure 16. Fluorescent intercalator displacement (FID) assay. The method is suitable for screening a library to identify a single compound with affinity for a given DNA sequence, or, for screening DNA libraries to establish the sequence preference of a DNA binding agent.

taining the sequence of interest, or a library of hairpin deoxyoligonucleotides containing all four bp (136 hairpins) or five bp (512 hairpins) sequences individually displayed in 96-well plates are treated with the intercalator, yielding a fluorescence increase upon DNA binding.[360] Addition of a DNA-binding compound results in a decrease in fluorescence owing to displacement of the bound intercalator. The decrease in fluorescence is directly related to the extent of DNA binding, which provides relative DNA binding affinities and a rank-order binding. This rank-order binding profile can be used in the assay of a library of compounds against a single sequence to identify the best candidate ligands for a targeted sequence. Alternatively, it can entail the assay of a single compound against a library of hairpin DNAs containing all possible sequences thereby establishing a high-resolution profile of its DNA-binding selectivity.

5.2.1. *Assaying Libraries of Compounds Against a Target Sequence Exemplified with the Discovery of Inhibitors of Androgen Receptor-Mediated Gene Transcription*

The FID assay can be used to rapidly screen a library of compounds against any target sequence or ensemble of sequences in a 96-well format. This approach was utilized to identify potent and selective inhibitors of androgen receptor-mediated gene transcription (Figure 17). The emergence of hormone-independent, constituently active androgen receptor dimer, which is unresponsive to competitive antagonist treatment, is responsible for prostate-cancer relapse that is resistant to chemotherapeutic treatment. A potentially effective treatment for such resistant prostate cancer could entail administration of a DNA-binding agent selective for the androgen response element (ARE)-consensus sequences. Such a treatment would competitively inhibit the constitutively active androgen receptor DNA binding and its transcription activation.

We used the FID assay with two hairpin deoxyoligonucleotides containing two related sequences of the androgen response element (ARE), the 14-bp ARE-consensus and PSA-ARE-3 sequences. The latter sequence contains a five-bp AT-rich sequence that could be anticipated to bind distamycin-like structures while the former contains a similar sequence interrupted by a GC bp. In the screening of our library, several candidates were identified which bound the AT-rich site of the PSA-ARE-3 sequence. One analogue (**49**, Figure 17) maintained the high-affinity binding to the ARE-consensus sequence. Compound **49** appeared suitable as an inhibition of the gene transcription initiated by hormone insensitive androgen-receptor dimerization and DNA binding characteristic of therapeutic-resistant prostate cancer.[87] Notably, **49** exhibited potent ($IC_{50} = 8$ nM) and selective (ca. 40-fold) inhibition of androgen receptor-mediated gene transcription in a cell-based reporter assay, albeit requiring liposome delivery of the compound for cell penetration and observation of the activity.

5.2.2. *High-Resolution Definition of Sequence Selectivity*

This FID assay is especially powerful for establishing the complete sequence selectivity of a DNA-binding agent. The technique entails the measurement of the loss of fluorescence derived from the displacement of ethidium bromide from hairpin deoxyoligonucleotides containing all possible five (512 hairpins) or four-bp sequences (136 hairpins) displayed in a 96-well format (Figure 18). In selected instances, the change in fluorescence of the DNA-binding compound itself can be monitored. This includes characterization of modest and nonbinding sequences not accessible by other techniques (i.e., footprinting), and such binding profiles may prove instrumental in distinguishing candidate ligands that display selectivity for an ensemble of related sequences characteristic of the consensus sequences of transcription factors. Technical issues associated with the use of the assay were investigated, multiple compounds of varying affinities were examined, and its use in comparing the DNA binding properties of distamycin A, netropsin, DAPI, Hoechst 33258, and berenil were

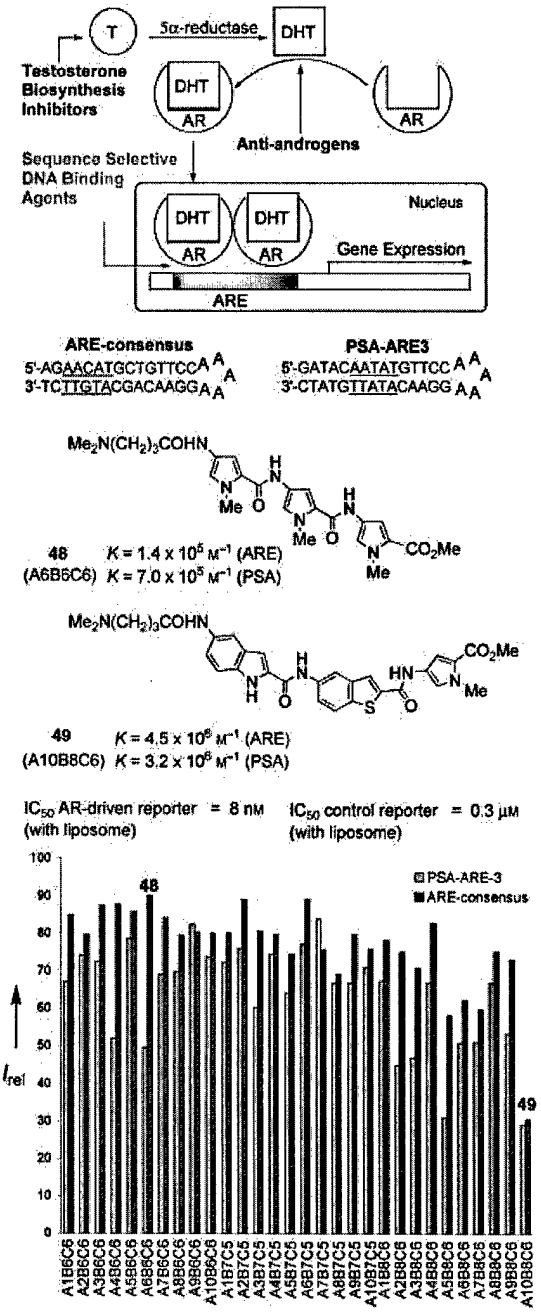


Figure 17. FID assay used in the screening for inhibitors of androgen receptor (AR) mediated gene transcription by evaluating the affinity to PSA-ARE3 and ARE-consensus sequences.

disclosed.[97] The cost of the 512 hairpins for a single assay is about $100 (ethidium bromide) or less (for thiazole orange).[98] The technique is nondestructive indicating that immobilization of the hairpins (chips, beads, glass slides) could permit their rinsing and reuse in subsequent assays. This hairpin immobilization and reuse would remove the barrier to

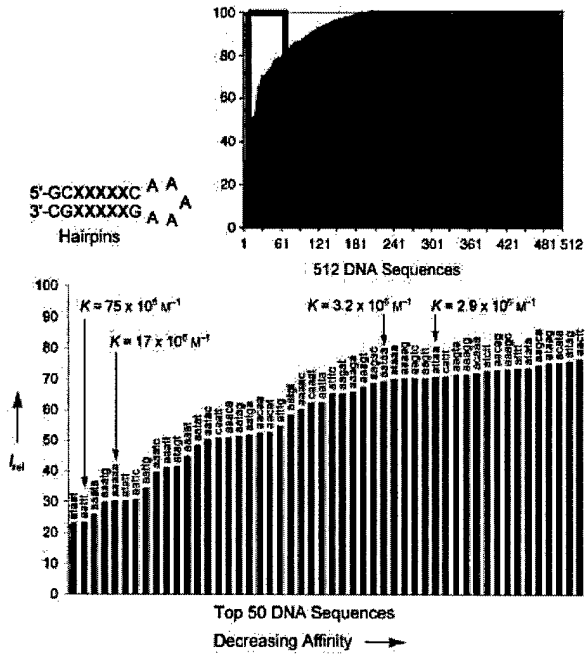


Figure 18. Determination of distamycin A sequence selectivity in one experiment by using the FID assay. The relative affinity for a given DNA sequence is proportional to the decrease in fluorescence. The affinity decreases with increasing AT content. The absolute binding constants obtained are comparable to those obtained with other techniques.

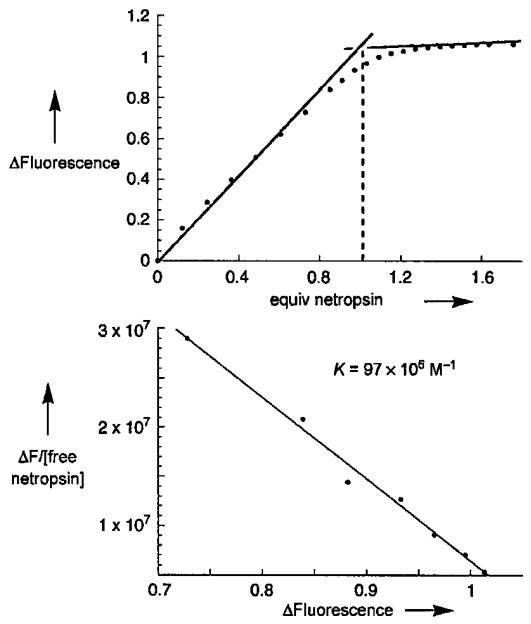


Figure 19. Top: Titration of netropsin versus the hairpin containing 5'-AATTT-3' at 1.1 μM (8.8 μM/bp), the intersection is at 1.05 equivalents which indicates a 1:1 binding and bottom: Scatchard plot (Slope = $-K$).

examining libraries of longer sequences (> 5 bp) and extend the use of the technology to binding site sizes typical of proteins.

5.2.3. *Scatchard Analysis of Titration Curves: Absolute Binding Constants and Stoichiometry of Binding*

The FID assay may be used to establish binding constants for any given sequence through quantitative titration. Several methods of establishing binding constants were examined and the most reliable is a Scatchard analysis of the titration binding curve which also determines the stoichiometry of binding (Figure 19).[97] Direct measurement of binding constants of fluorescent DNA-binding compounds (DAPI, Hoechst 33258) compared favorably to those derived from the indirect displacement of ethidium bromide and thiazole orange (Table 5). As a demonstration of the power of this technique, the absolute binding constants for netropsin with each four-bp AT sequence were determined establishing its full relative and absolute selectivity (Table 6).[98]

Table 5: Binding constants of DAPI and Hoechst 33258.

A. Ethidium bromide and thiazole orange binding constants		
Hairpin	K_{EB} [$\times 10^6$ M^{-1}][a]	K_{TO} [$\times 10^6$ M^{-1}][a]
5'-AAAAA-3'	0.27 (4.5)[b]	1.2 (3.5)[b]
5'-AATTT-3'	0.42 (3.6)	1.3 (3.3)
5'-ATTAA-3'	0.87 (3.3)	1.0 (3.3)
5'-AATAA-3'	0.73 (3.3)	1.1 (3.3)

DAPI binding constants						
DNA Sequence	K [$\times 10^6$ M^{-1}][c]		K [$\times 10^6$ M^{-1}][d]		K [$\times 10^6$ M^{-1}][e]	
5'-AATTT-3'	58	(0.91)	110	(0.94)	120	(1.1)
5'-AATAA-3'	26	(0.83)	59	(0.98)	87	(0.94)
5'-ATTAA-3'	24	(0.83)	52	(0.88)	77	(1.0)
5'-AAAAA-3'	n.d.[f]	n.d.[f]	50	(0.95)	65	(0.99)

Hoechst 33258 binding constants						
DNA Sequence	K [$\times 10^6$ M^{-1}][c]		K [$\times 10^6$ M^{-1}][d]		K [$\times 10^6$ M^{-1}][e]	
5'-AATTT-3'	75	(0.90)	88	(0.96)	177	(0.98)
5'-AAAAA-3'	36	(0.86)	72	(0.98)	145	(0.92)
5'-AATAA-3'	28	(0.77)	34	(1.06)	83	(1.01)
5'-ATTAA-3'	13	(0.89)	19	(0.91)	47	(1.05)

[a] Established for the hairpin deoxyoligonucleotides containing the indicated sequence by direct titration. [b] Stoichiometry of binding established experimentally. [c] Thiazole orange displacement (experimental stoichiometry of binding). [d] Ethidium bromide displacement (experimental stoichiometry of binding). [e] Direct fluorescence titration (experimental stoichiometry of binding). [f] n.d. = not determined.

5.2.4. *Rapid Characterization of DNA Binding Modes*

The FID assay has been utilized to distinguish between and establish binding modes and binding-site sizes. This work has permitted the first characterization of a novel substituted β-alanine hairpin polyamide and the cooperative 2:1 side-by-side parallel (versus antiparallel) binding of iminodiacetic acid linked pyrrole polyamides. Polyamides composed of *N*-methylpyrrole, *N*-methylimidazole, and a growing set of structural analogues bind in the DNA minor groove with predictable sequence selectivity and high affinity.[350,361–363] The head-to-tail linkage of such polyamides with the five-atom linker γ-aminobutyric acid (γ) has been shown to

Table 6: Comparative binding constants.

Titrant	DNA Sequence	K [$\times 10^6$ M^{-1}]	Binding stoichiometry
ethidium bromide	5′-AATT	2.7	2.92
thiazole orange	5′-AATT	15	3.05
Titrant	DNA Sequence	Ethidium bromide K [$\times 10^6$ M^{-1}]	Thiazole orange K [$\times 10^6$ M^{-1}]
Netropsin	5′-AAAT	127	113
	5′-AAAA	92	71
	5′-AATT	65	54
	5′-AATA	64	44
	5′-ATTA	45	35
	5′-ATAT	41	33
	5′-ATAA	34	18
	5′-TAAA	26	15
	5′-TTAA	11	8
	5′-TATA	11	8

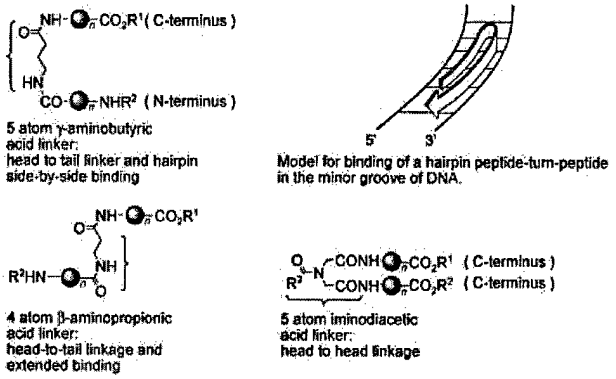


Figure 20. Linked polyamides for DNA binding.

provide hairpin polyamides that mimic the 2:1 side-by-side antiparallel binding of the unlinked polyamides (Figure 20), enhance the binding affinity 100 to 10000-fold, and improve the binding selectivity.[364] In contrast, polyamides incorporating a one-carbon-atom shorter head-to-tail linker, β-alanine (β), bind preferentially in an extended conformation forming 1:1 or side-by-side antiparallel 2:1 complexes.[364,365] By using the FID assay and a series of DNA hairpins containing a AT-rich binding site, the size of which was systematically varied, the stoichiometry of binding (1:1 vs 2:1) and the binding mode (hairpin vs extended) for such distamycin-derived polyamides can be rapidly and easily established (Figure 20).[83,84]

Using this approach, appropriate substitution of a β-alanine linker was established to induce hairpin instead of extended binding providing an alternative hairpin linker to γ (Figure 21).[83,84] These polyamide analogues were obtained using a parallel solution-phase approach involving a series of EDCI-mediated coupling reactions (Scheme 12).

As a natural extension of these studies, the behavior of head-to-head iminodiacetic acid (IDA) linked polyamides was characterized. Not only was IDA the template on which many other libraries had been generated (see Section 3), but it constituted a unique five-atom linker bearing an additional functionalization site at the central nitrogen atom of its structure, and represented an unusual head-to-head as opposed to head-to-tail linkage of the polyamides. The candidate IDA-linked pyrrole polyamides were prepared with solution-phase synthesis methodology (Scheme 13) and found to bind in a unique manner exhibiting cooperative extended 2:1 binding across a 9–10-bp site in what appears to be the first well-characterized example of parallel, side-by-side binding of such polyamides (Figure 22).[84]

5.3. LEF-1/β-Catenin

A second class of inhibitors of LEF-1/β-catenin mediated gene transcription (see Section 4.4 for the first class) were derived from the libraries of DNA minor-groove-binding ligands.[87] The unpublished results of the screening of one such library are shown in Figure 23 resulting in the identification of two closely related lead structures **69** and **70** that inhibit TOPFLASH transcription 10-fold, exhibit a ≧threefold TOPFLASH-FOPFLASH selectivity, inhibit LEF-1 DNA binding in a gel-shift assay, and inhibit the β-catenin induced transformation of the colon cancer cell line SW480. Their DNA-binding selectivities were established in our FID assay.[87] This constitutes only the second example of a small-molecule inhibitor of LEF-1/DNA binding (protein–DNA interaction) leading to inhibition of gene transcription, the first being a polyamide from Dervan et al.[366,367]

6. Natural Product Libraries

Natural product leads for existing or new targets often emerge as superb starting points for drug development. The merits of enlisting a natural product lead have been chronicled in many reviews and monographs,[368] and may be attributed to several features. They have emerged from a biological milieu in which they express their properties ensuring that they embody many drug-like qualities. No doubt, it is the integration of many functions into a single compact structure that distinguishes natural product leads from those created in the laboratory. As the exploration of the properties of complex natural products becomes increasingly more sophisticated with the technological advances being made in their screening and evaluation and as structural details of their interaction with biological targets becomes more accessible, the opportunities for providing unique solutions to complex biological problems have grown. A powerful complement to the examination of the naturally derived agents themselves is the preparation and subsequent examination of key partial structures, and agents containing deep-seated structural modifications. Well conceived deep-seated structural modifications may be used to address the structural basis of the natural product's interactions with biological targets and the define fundamental relationships between structure, functional reactivity, and properties. In these studies, the challenging problem is to understand the beautiful solutions and subtle design elements that nature has provided in the form of a natural product and work to extend

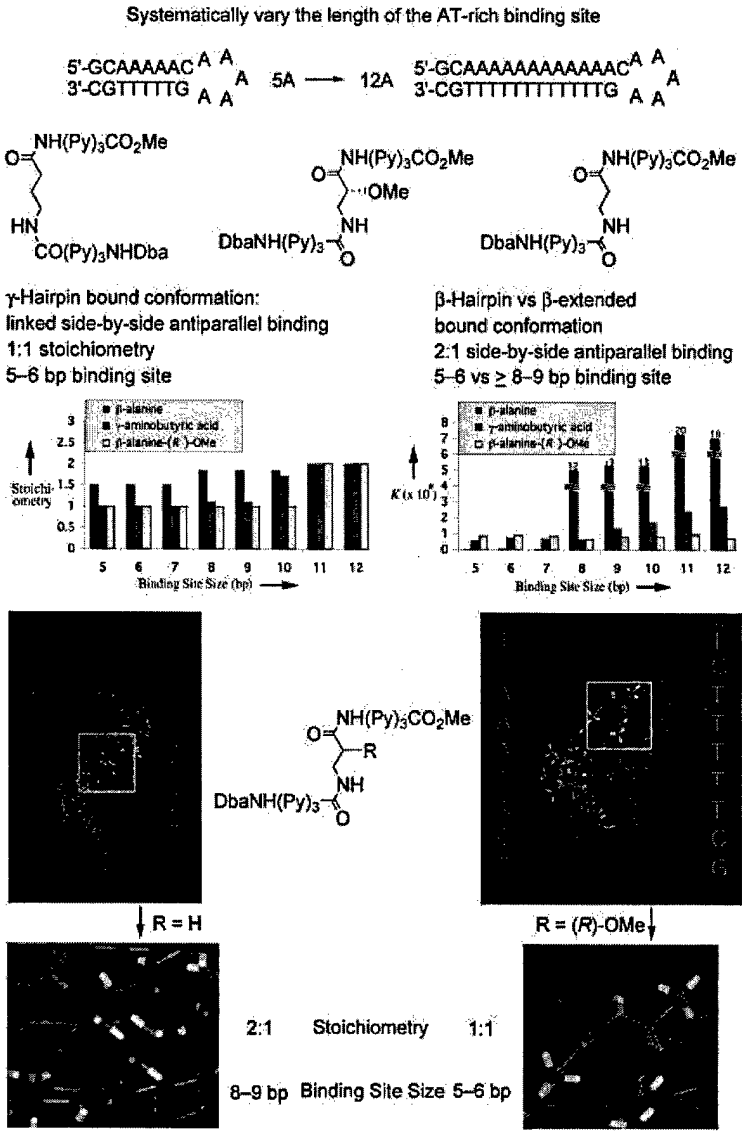


Figure 21. Hairpin binding of a substituted β-alanine linked polyamide. Bottom left: β-alanine extended antiparallel 2:1 side-by-side binding. Bottom right: $\beta^{(R)\text{-OMe}}$-alanine: antiparallel 1:1 side-by-side hairpin binding.

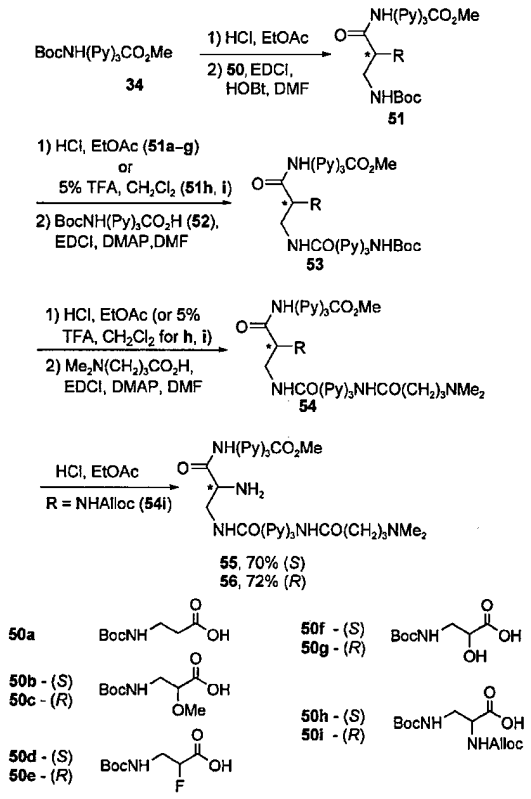
the solution through rational design elements to provide more selective, more efficacious, or more potent agents designed specifically for the problem or target under investigation.

Still today, nine of top-20 best-selling drugs were derived or developed from natural product leads with combined sales of > $16 billion (1999).[407] Consequently, natural products are arguably one of the most successful approaches to modern drug discovery, and it is not surprising that combinatorial methods have been evolved to exploit the importance of such leads (Table 7).[369–407] The solution-phase synthesis approach we developed has proven successful for the generation of libraries of selected natural products. In addition to the synthesis of 2640 analogues of distamycin (Section 5.1.1), 132 analogues of CC-1065 and the duocarmycins (Section 5.1.2), and a prototype analogue of triostin A (Section 5.1.3), the approach was used to create a library of HUN-7293 analogues exploring each structural detail of the natural product. This example is one in which the biological target for the natural product has not been established. Even in the absence of a known target, a functional screen and a library of analogues was used to define the pharmacophore of the natural product.

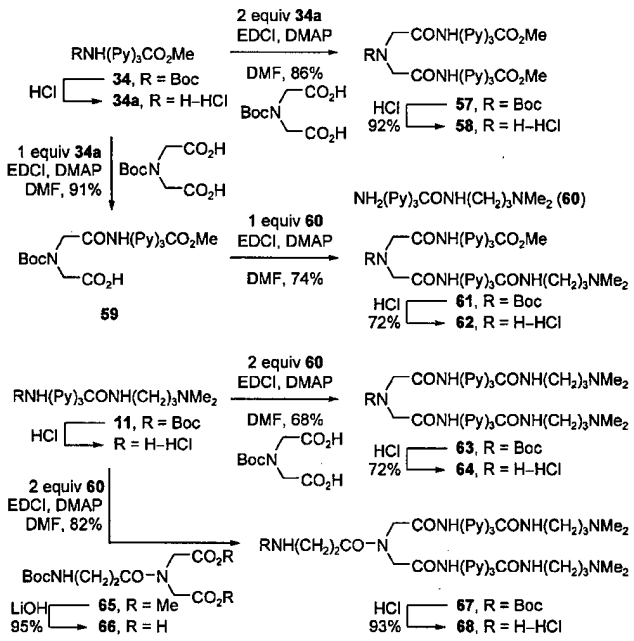
6.1. HUN-7293: Single-Point Changes Introduced by Total Synthesis of a Library of Analogues: A "Chemical" Mutagenesis Approach to Defining Structure–Function Relationships

The naturally occurring cyclic heptadepsipeptide HUN-7293 (Scheme 14) was first isolated in 1992 from a fungal broth in a screen for inhibitors of inducible cell adhesion molecule expression.[408] Its structure was subsequently established by ^{1}H NMR spectroscopic methods and confirmed by X-ray analysis.[408] Independently, the same cyclic depsipeptide was isolated by a Japanese group from a different fungal species based on a screen for anti-HIV compounds.[409] Molecules that bind with endothelial cells, which include intercellular adhesion molecule 1 (ICAM-1), vascular cell adhesion molecule 1 (VCAM-1), and E-selectin play important roles in the immune response by recruiting and regulating leukocyte migration and cell-to-cell interactions at a site of inflammation. Modulation of these interactions through the inhibition of such cell adhesion molecule expression should prove useful for the treatment of chronic inflammatory and autoimmune diseases characterized by their overexpression or upregulation. As detailed in recent studies,[410] diverse signals act on endothelial cells to activate members of the nuclear factor of the Kappa B (NF-κB) transcription factor family. NF-κB family members are cytoplasmic homo- or heterodimer proteins inactive when complexed with members of a family of inhibitor proteins, IκB. Phosphorylation of IκB, releasing the active NF-κB, results in translocation to the nucleus and initiation of transcription of its targeted genes including the cell surface adhesion molecules. Few other groups have detailed work on small-molecule inhibitors of the expression[413–415] of cell surface adhesion molecules, aside from antisense[411] and antibody[412] approaches, and unlike HUN-7293, each target the NF-κB signal transduction pathway.

The first total synthesis of HUN-7293 was disclosed using a convergent approach which served to define several subtle elements essential to its preparation including backbone ester formation by a Mitsunobu displacement and the identification of MLEU^3-LEU^4 as an effective macrocyclization site (Scheme 14).[74] The subsequent implementation of this



Scheme 12. The parallel synthesis of β-alanine linked pyrrole polyamides. Py = *N*-methylpyrrole, Im = *N*-methylimidazole, HOBt = 1-hydroxybenzotriazole, TFA = trifluoroacetic acid, Alloc = allyloxycarbonyl.



Scheme 13. Synthesis of pyrrole polyamides linked with an IDA scaffold.

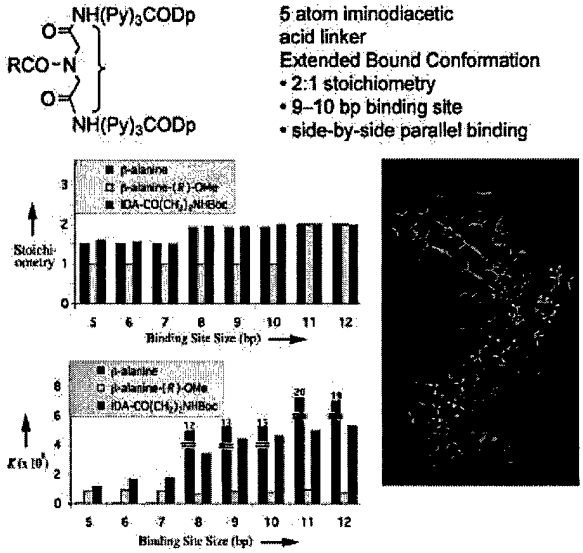


Figure 22. Extended 2:1 parallel side-by-side binding of IDA-linked polyamides. The best results are obtained with R = $(CH_2)_2$NHBoc.

approach in the convergent parallel synthesis of a library of key analogues of HUN-7293 was used to define in detail the structure–function relationships of the natural product (Scheme 15).[75,76]

A solution-phase approach to the simultaneous preparation of the library was utilized enlisting simple acid–base liquid–liquid extractions for isolation and purification of the synthetic intermediates and final products (Scheme 16). Notably, this approach provided each intermediate and final product sufficiently pure for subsequent use even in the multistep synthesis of HUN-7293 and its structurally challenging analogues. The library included an alanine scan of the seven residues, an *N*-methyl deletion of the three *N*-methyl residues, and fundamental simplifications in each of the nonstandard amino acid side chains allowing the identification of key residues and structure features contributing to the biological properties. The simultaneous inclusion of simplified analogues of the nonstandard amino acid side chains permitted the identification of not only essential structural features, but also non-obvious and subtle structural features.

Alanine-scanning mutagenesis is an informative method for defining epitopes in proteins.[416] By replacing each amino acid side chain with a methyl group while retaining the stereochemistry of the peptide backbone, the significance of each individual side chain can be identified. Although this approach has its limitations, it has been used to provide a general insight into structure–function relationships of a protein, especially when the functional mechanism is unknown or its 3D structure is unavailable.[416] Similarly, in the absence of a defined target, the importance of each residue of HUN-7293 could not be anticipated. Along with additional modifications and simplifications in each of the individual side chains, an alanine scan of the structure was conducted. The results of their examination are sum-

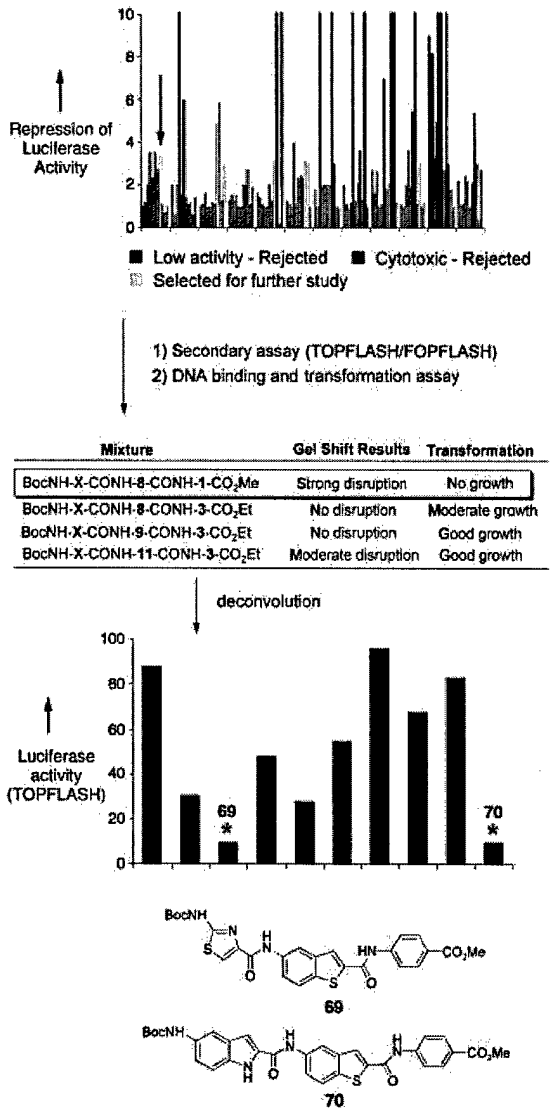


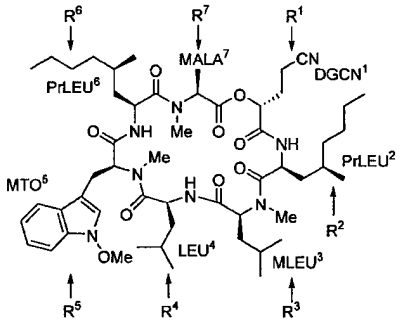
Figure 23. Disruption of LEF-1 DNA binding and inhibition of LEF-1/β-catenin-mediated gene transcription. The lead structures **69** and **70** have a tenfold transcription repression (TOPFLASH) and threefold selectivity (TOPFLASH/FOPFLASH)

marized in Table 8 and they provided an initial, albeit surprisingly incomplete, overview of the relative importance of each residue. With the exception of residue 7, which already contains a methyl group in the natural product (R^7 = Me), the replacement of each side chain with a methyl group led to a reduction in potency. However, the magnitude of this loss in activity varied over a 1000-fold range with the R^1 replacement being only about two-times lower, the R^2–R^4 replacements leading to 10–100-fold reductions, and the R^5 and R^6 replacements leading to even larger, over 100–1000-fold, reductions with R^5 being by far the most significant residue. Just as interesting and despite the potency reductions, the methyl substitutions for the R^1–R^4 side chains enhanced the VCAM-1 versus ICAM-1 selectivity indicating that the changes in this half of the molecule affected inhibition of ICAM-1 expression more significantly, whereas this selectivity was lost along with the activity with methyl substitutions at R^5 and R^6 with the change at R^5 again being especially detrimental. From these results alone, one could anticipate that substantial changes in the R^1 side chain could be tolerated including efforts to enhance selectivity or physiochemical, distribution, stability, and toxicity profiles, whereas changes to R^2–R^4 might prove more challenging, and those to R^5 (a complete set of modifications of R^5 is

Table 7: Natural product libraries.

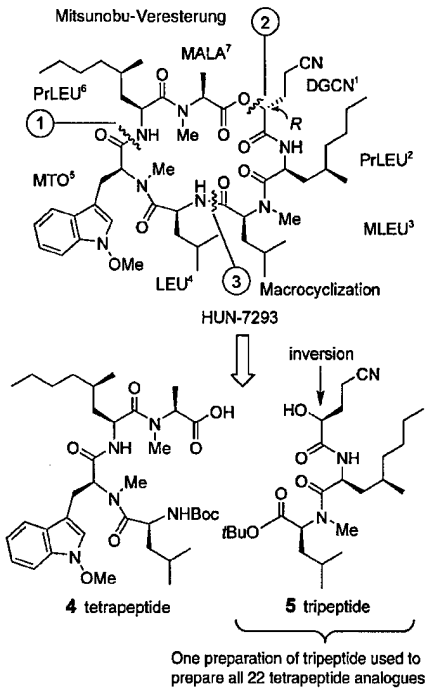
Natural product	Library format	Size	Reference
Lavendustin A	Solid Phase	61	[370]
Balanol	Solid Phase	32	[371]
Olomoucine	Solution and Solid Phase	364	[372]
Rauwolfa Alkaloids	Solid Phase	792	[373]
Epothilones	Solid Phase	180	[374]
Prostaglandin E_2 and $F_{2\alpha}$	Soluble Polymer Support	2/16/6	[375]
Taxoids	Solid Phase	400	[376]
Oscillamide Y	Solid Phase	4	[377]
(*S*)-Zearalenone	Solid Phase	1	[378]
Cycloserine	Solid Phase	80	[379]
Flavones	Solution Phase	36	[380]
Muscone	Solid Phase	12	[381]
Sarcodictyin	Solution Phase	15/53	[382]
Taxoids	Solution Phase	26	[383]
Prostaglandin E and F	Solid Phase	26	[384]
Fumitremorgin C	Solid Phase	42	[385]
Indolactam V	Solid Phase	31	[386]
Vitamin D_3	Solid Phase	1	[387]
Kramerixin	Solution Phase	120	[388]
Vancomycin	Solution and Solid Phase	39,304	[389]
HUN-7293	Solution Phase	1/40	[75,76]
Carpanone	Solid Phase	6	[390]
Distamycin A	Solution Phase	2640	[87]
Distamycin A	Positional Scanning	1000	[86]
Triostin A	Solution Phase	1	[72]
Fumiquinazoline	Solid Phase	27	[391]
Benzopyrans	Solution and Solid Phase	>10000	[392]
Steroids	Solid Phase	20	[393]
Curacin A	Solution Phase	18	[394]
Stipiamide	Solution Phase	42	[395]
CC-1065	Solution Phase	132	[88]
Galanthamine	Solid Phase	2527	[397]
Mappicine	Solution Phase	112/560	[396]
Tambjamine	Solution Phase	10	[398]
Vitamin D_3	Solid Phase	72	[399]
Vitamin D_3	Solution and Solid Phase	5	[400]
Psammaplin A	Solution Phase	3828	[401]
Hapalosin	Solution and Solid Phase	6	[402]
Nakijiquinones	Solution Phase	56	[403]
Steroids	Solution Phase	298	[404]
Ceramide	Solid Phase	528	[405]
Dysidolide	Solid Phase	8	[406]
Natural product-like	–	–	[369] (review)

Table 8: Alanine scan of HUN-7293. Inhibitory activity (IC_{50} [nM]) for VCAM-1 and ICAM-1 protein expression in human microvascular cell line (HMEC-1).[a]



Compound	VCAM-1	ICAM-1	Selectivity	Rel. potency
HUN-7293[b]	1	24	24	1
79, R^1 = Me	2.3 ± 0.5	178 ± 18	77	0.4
80, R^2 = Me	100 ± 34	6300 ± 370	63	0.01
81, R^3 = Me	56 ± 11	4100 ± 2500	73	0.02
82, R^4 = Me	16 ± 1.4	700 ± 20	43	0.06
83, R^5 = Me	1800 ± 10	2800 ± 210	1.6	0.0006
84, R^6 = Me	270 ± 40	1400 ± 300	5.2	0.004

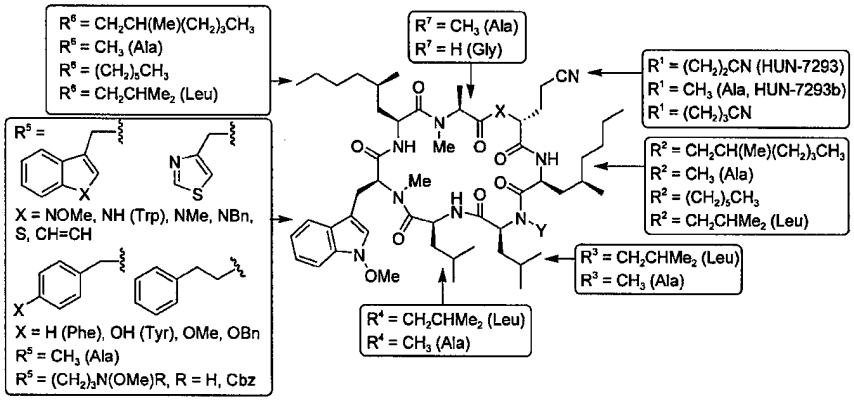
[a] Unless otherwise indicated, the substituents are those found in HUN-7293. [b] R^7 = Me, HUN-7293.



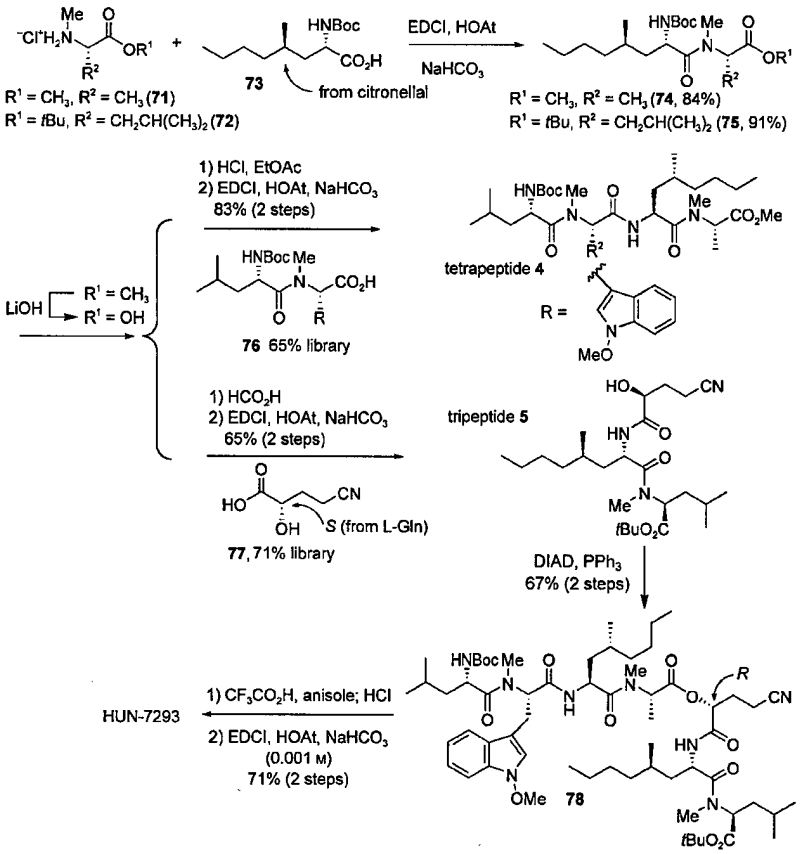
Scheme 14. Parallel synthesis of a library of HUN-7293 analogues: divergent preparation of a library by using a convergent total synthesis. All the intermediates and products were isolated and purified by liquid–liquid acid–base extraction. ①–③ indicate disconnections.

presented in Table 9) and R^6 are likely to be especially detrimental. Consequently, the results of the more subtle side-chain modifications proved especially valuable at addressing inaccuracies in interpreting such an incomplete data set from the alanine scan.

An assessment of the impact of each residue, each side chain, and each structural feature of HUN-7293 was established enlisting a complete library of analogues composed of systematic single-point changes in the structure prepared by parallel solution-phase synthesis. Over 40 new compounds were assembled by superimposing a divergent synthetic strategy onto a convergent total synthesis. The combination of the beneficial attributes of a conventional convergent synthetic strategy with the divergent preparation of a series of structurally related analogues can only be conducted using solution-phase, not solid-phase, techniques and avoids the more repetitive independent linear synthesis of each analogue. This generalized library approach of chemical mutagenesis of the cyclic depsipeptide provided a detailed first level structure–activity study that subsequent optimization efforts can be confidently based on. One of the most striking conclusions of the study was that a simple alanine scan of the naturally occurring depsipeptide may provide a sense of the



Scheme 15. Construction of HUN-7293 pharmacophore library via solution-phase parallel synthesis; X = O, NH; Y = H, Me; Bn = benzyl, Cbz = benzyloxycarbonyl. For more details see text.



Scheme 16. Synthesis of HUN-7293 and analogues. A Mitsunobu esterification permits the use of the L-Gln-derived alcohol precursor **77** and avoids the racemization observed with activation of the NMe-Ala C-terminus. With **78** epimerization upon macrolactamization is minimized with DPPA (diphenylphosphorylazide). HOAt = 1-hydroxy-7-azabenzotriazole, DIAD = diisopropyl azodicarboxylate.

importance of each residue, but it provides an incomplete data set that is problematic to draw substantive conclusions from since each nonstandard residue or side chain, including residues 5 and 6, could be simplified without substantially influencing the potency. Although similar approaches have been disclosed to optimize or explore a single residue within a naturally occurring cyclic peptide, we are not aware of an instance where all the residues were examined simultaneously. Moreover, a more common approach to initial structure–activity relationship studies of such natural products has been systematic derivatization or degradation studies to establish the structural features important for activity. These efforts, like an alanine scan, provide an incomplete data set on which to base further studies.

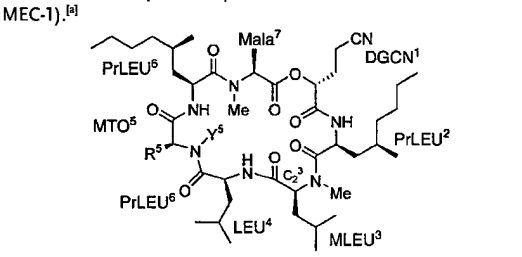
7. Summary and Outlook

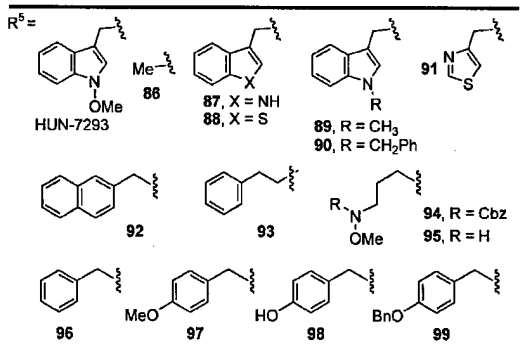
The complementary techniques of solid-phase and solution-phase combinatorial chemistry provide rare opportunities to rapidly probe challenging targets, especially those for which no known small-molecule leads exist and for which there is little or no structural information available. Split-and-mix solid-phase synthesis and solution-phase mixture synthesis have substantially increased the capacity for preparing candidate libraries and their value has increased as deconvolution techniques for rapidly or easily identifying screening leads have been introduced. Among its many attributes, solution-phase library synthesis permits the utilization of the widely accepted convergent synthesis strategy, easily accommodates mixture synthesis, is adaptable to dynamic- or selection-screening including target-assembled ligands, and is generally not limited by scale. Moreover, the advances in soluble polymer-supported synthesis and the widespread success of strategies enlisting polymer-supported reagents, scavengers, and workup reagents (inverse solid-phase synthesis)[165] suggest that the common assumption that library synthesis is best conducted on solid phase may not always be accurate and that the value of solution-phase library synthesis may still be underestimated.

Our present library of approximately 40000 compounds is in the process of being expanded to 1000000 compounds with a focus on assembling a series of libraries capable of adequately, if not fully, exploring pharmacophore space in a single round of screening, incorporating specific design elements for targeting protein–protein interactions (e.g. RGD mimetics, α-helix and β-sheet mimetics), exploring new templates, and implementing new solution-phase techniques including target-assisted synthesis for selection of candidate ligands. The fact that the libraries are generated on a scale permitting their repeated use in new assays as the opportunities arise ensures their value will only increase over time.

It is likely that the years ahead will be characterized by an increasing focus on the application of the technology more than the methodology itself, and this may very well be the era in which the most useful approaches naturally emerge. It will also mark a time where the problem selection may distinguish the notable contributions from those that are overlooked. The future issue may well be not just the identification of effective modulators (inhibitors or mimics) of target protein–protein or protein–DNA interactions, but the identification of selective modulators. A first step in such an analysis for protein–DNA interactions was taken with the introduction of the FID assay which permits the simultaneous examination of all possible DNA sequences. This allows an assessment of affinity for not

Table 9: R^5 modification of HUN-7293. Inhibitory activity (IC_{50} [nM]) for VCAM-1 and ICAM-1 protein expression in human microvascular cell line (HMEC-1).[a]





Compound	VCAM-1	ICAM-1	Selectivity	Rel. potency
HUN-7293	1	24	24	1
85, Y^5 = H	> 10000	> 10000	n.d.[b]	< 0.0001
86	1800 ± 10	2800 ± 210	1.6	0.0006
87	44 ± 11	510 ± 140	12	0.02
88	1.7 ± 1.6	19 ± 1.2	11	0.6
89	21 ± 1.6	186 ± 3	9	0.05
90	19 ± 1.2	16 ± 13	0.8	0.05
91	38 ± 5	2100 ± 170	55	0.03
92	11 ± 2.6	71 ± 35	6	0.09
93	29 ± 6	94 ± 13	3	0.03
94	380 ± 120	890 ± 340	2.3	0.003
95	1400 ± 230	> 7000	> 5	0.0007
96	2.3 ± 0.6	43 ± 6	19	0.4
97	63 ± 4	720 ± 90	11	0.02
C_2^3-epi-**97**	190 ± 20	9700 ± 3300	51	0.005
98	810 ± 130	8200 ± 120	10	0.001
99	280 ± 7	480 ± 65	1.7	0.004
C_2^3-epi-**99**	1100 ± 240	2600 ± 440	2.4	0.0009

[a] Y^5 = Me (HUN-7293 unless otherwise stated). [b] n.d. = not determined.

only the desired target (sequence), but also a simultaneous assessment of selectivity among all targets (sequences) that may be used to distinguish between otherwise similar leads. Proteomics approaches towards examining ligand interactions with all proteins are emerging[417] that may be used to establish ligand selectivity,[99] and analogous approaches to simultaneous or parallel screening for modulating protein–protein interactions[418] are on the horizon.

We gratefully acknowledge the financial support of the National Institutes of Health (CA78045, CA41986, CA41101, EY14174), the Department of Army (DAMD17-02-1-0135), the Skaggs Institute for Chemical Biology, and predoctoral fellowship for K.C. (Thomas and Sharon Bauce Foundation Fellowship). K.C. is a Skaggs Fellow.

Received: January 24, 2003 [A574]

[1] M. R. Pavia, T. K. Sawyer, W. H. Moos, *Bioorg. Med. Chem. Lett.* **1993**, *3*, 387–396.
[2] M. A. Gallop, R. W. Barrett, W. J. Dower, S. P. A. Fodor, E. M. Gordon, *J. Med. Chem.* **1994**, *37*, 1233–1251.
[3] E. M. Gordon, R. W. Barrett, W. J. Dower, S. P. A. Fodor, M. A. Gallop, *J. Med. Chem.* **1994**, *37*, 1385–1401.
[4] N. K. Terrett, M. Gardner, D. W. Gordon, R. J. Kobylecki, J. Steele, *Tetrahedron* **1995**, *51*, 8135–8173.
[5] J. S. Früchtel, G. Jung, *Angew. Chem.* **1996**, *108*, 19–46; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 17–42.
[6] J. A. Ellman, *Acc. Chem. Res.* **1996**, *29*, 132–143.
[7] E. M. Gordon, M. A. Gallop, D. V. Patel, *Acc. Chem. Res.* **1996**, *29*, 144–154.
[8] L. A. Thompson, J. A. Ellman, *Chem. Rev.* **1996**, *96*, 555–600.
[9] F. Balkenhohl, C. von dem Bussche-Hünnefeld, A. Lansky, C. Zechel, *Angew. Chem.* **1996**, *108*, 2436–2487; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 2288–2337.
[10] I. M. Chaiken, K. D. Janda, *Molecular Diversity and Combinatorial Chemistry: Libraries and Drug Discovery*, American Chemical Society, Washington, DC, **1996**.
[11] G. Jung, *Combinatorial Peptide and Nonpeptide Libraries: a Handbook*, VCH, Weinheim, **1996**.
[12] A. W. Czarnik, S. H. DeWitt, *A Practical Guide to Combinatorial Chemistry*, American Chemical Society, Washington, DC, **1997**.
[13] A. Nefzi, J. M. Ostresh, R. A. Houghten, *Chem. Rev.* **1997**, *97*, 449–472.
[14] N. K. Terrett, *Combinatorial Chemistry*, Oxford University Press, Oxford, UK, **1998**.
[15] D. Obrecht, J. M. Villalgordo, *Solid-Supported Combinatorial and Parallel Synthesis of Small-Molecular-Weight Compound Libraries*, Pergamon/Elsevier, Oxford, UK, **1998**.
[16] B. Agnew, R. F. Sevice, M. Enserink, D. Normile, E. Marshal, C. Sander, M. R. Reich, *Science* **2000**, *287*, 1951–1981.
[17] R. Storer, *Drug Discovery Today* **1996**, *1*, 248–254.
[18] A. Chucholowski, T. Masquelin, D. Obrecht, J. Stadlwieser, J. M. Villalgordo, *Chimia* **1996**, *50*, 525–530.
[19] D. M. Coe, R. Storer, *Annu. Rep. Comb. Chem. Mol. Diversity* **1997**, *1*, 50–58.
[20] N. Bailey, A. W. J. Cooper, M. J. Deal, A. W. Dean, A. L. Gore, M. C. Hawes, D. B. Judd, A. T. Merritt, R. Storer, S. Travers, S. P. Watson, *Chimia* **1997**, *51*, 832–837.
[21] L. M. Gayo, *Biotechnol. Bioeng.* **1998**, *61*, 95–106.
[22] H. An, P. D. Cook, *Recent Res. Dev. Org. Chem.* **1998**, *2*, 473–488.
[23] R. Ferritto, P. Seneci, *Drugs Future* **1998**, *23*, 643–654.
[24] D. M. Coe, R. Storer, *Annu. Rep. Comb. Chem. Mol. Diversity* **1999**, *2*, 1–8.
[25] D. M. Coe, R. Storer, *Mol. Diversity* **1999**, *4*, 31–38.
[26] V. Austel, *Comb. Chem.* **1999**, 77–123.
[27] M. J. Suto, *Curr. Opin. Drug Discovery Dev.* **1999**, *2*, 377–384.
[28] A. Ganesan, *Drug Discovery Today* **2002**, *7*, 47–55.
[29] H. An, *Front. Biotechnol. Pharm.* **2000**, *1*, 83–113.
[30] C. M. Baldino, *J. Comb. Chem.* **2000**, *2*, 89–103.
[31] H. An, P. D. Cook, *Chem. Rev.* **2000**, *100*, 3311–3340.
[32] O. Ramström, J.-M. Lehn, *Nat. Rev. Drug Discovery* **2002**, *1*, 26–36.
[33] S. Otto, R. L. E. Furlan, J. K. M. Sanders, *Curr. Opin. Chem. Biol.* **2002**, *6*, 321–327.
[34] H. Han, M. M. Wolfe, S. Brenner, K. D. Janda, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 6419–6423.

[35] D. J. Gravert, K. D. Janda, *Chem. Rev.* **1997**, *97*, 489–509.
[36] P. Wentworth, Jr., K. D. Janda, *Chem. Commun.* **1999**, 1917–1924.
[37] P. H. Toy, K. D. Janda, *Acc. Chem. Res.* **2000**, *33*, 546–554.
[38] N. A. Boyle, K. D. Janda, *Curr. Opin. Chem. Biol.* **2002**, *6*, 339–346.
[39] D. P. Curran, *Angew. Chem.* **1998**, *110*, 1230–1255; *Angew. Chem. Int. Ed.* **1998**, *37*, 1174–1196, Erratum: D. P. Curran, *Angew. Chem.* **1998**, *110*, 2569; *Angew. Chem. Int. Ed.* **1998**, *37*, 2292.
[40] D. P. Curran, *Med. Res. Rev.* **1999**, *19*, 432–438.
[41] J. Yoshida, K. Itami, *Chem. Rev.* **2002**, *102*, 3693–3716.
[42] R. A. Fecik, K. E. Frank, E. J. Gentry, L. A. Mitscher, M. Shibata, *Pure Appl. Chem.* **1999**, *71*, 559–564.
[43] H. Perrier, M. Labelle, *J. Org. Chem.* **1999**, *64*, 2110–2113.
[44] X. Wang, J. J. Parlow, J. A. Porco, Jr., *Org. Lett.* **2000**, *2*, 3509–3512.
[45] T. Bosanac, C. S. Wilcox, *J. Am. Chem. Soc.* **2002**, *124*, 4194–4195.
[46] S. W. Kaldor, M. G. Siegel, *Curr. Opin. Chem. Biol.* **1997**, *1*, 101–106.
[47] R. J. Booth, J. C. Hodges, *Acc. Chem. Res.* **1999**, *32*, 18–26.
[48] J. J. Parlow, R. V. Devraj, M. S. South, *Curr. Opin. Chem. Biol.* **1999**, *3*, 320–336.
[49] S. V. Ley, I. R. Baxendale, R. N. Bream, P. S. Jackson, A. G. Leach, D. A. Longbottom, M. Nesi, J. S. Scott, R. I. Storer, S. J. Taylor, *J. Chem. Soc. Perkin Trans. 1* **2000**, *23*, 3815–4195.
[50] S. V. Ley, I. R. Baxendale, *Nat. Rev. Drug Discovery* **2002**, *1*, 573–586.
[51] C. A. McNamara, M. J. Dixon, M. Bradley, *Chem. Rev.* **2002**, *102*, 3275–3300.
[52] D. E. Bergbreiter, *Chem. Rev.* **2002**, *102*, 3345–3384.
[53] T. J. Dickerson, N. N. Reed, K. D. Janda, *Chem. Rev.* **2002**, *102*, 3325–3344.
[54] D. L. Flynn, R. V. Devraj, J. J. Parlow, *Curr. Opin. Drug Discovery Dev.* **1998**, *1*, 41–50.
[55] D. L. Flynn, R. V. Devraj, W. Naing, J. J. Parlow, J. J. Weidner, S. Yang, *Med. Chem. Res.* **1998**, *8*, 219–243.
[56] D. L. Flynn, *Med. Res. Rev.* **1999**, *19*, 408–431.
[57] D. L. Boger, C. M. Tarby, L. H. Caporale, *J. Am. Chem. Soc.* **1996**, *118*, 2109–2110.
[58] S. Cheng, D. D. Comer, J. P. Williams, D. L. Boger, *J. Am. Chem. Soc.* **1996**, *118*, 2567–2573.
[59] S. Cheng, C. M. Tarby, D. D. Comer, J. P. Williams, L. H. Caporale, D. L. Boger, *Bioorg. Med. Chem.* **1996**, *4*, 727–737.
[60] D. L. Boger, W. Chai, R. S. Ozer, C.-M. Andersson, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 463–468.
[61] D. L. Boger, W. Chai, *Tetrahedron* **1998**, *54*, 3955–3970.
[62] D. L. Boger, J. Goldberg, C.-M. Andersson, *J. Org. Chem.* **1999**, *64*, 2422–2427.
[63] D. L. Boger, W. Jiang, J. Goldberg, *J. Org. Chem.* **1999**, *64*, 7094–7100.
[64] D. L. Boger, W. Chai, Q. Jin, *J. Am. Chem. Soc.* **1998**, *120*, 7220–7225.
[65] D. L. Boger, J. K. Lee, J. Goldberg, Q. Jin, *J. Org. Chem.* **2000**, *65*, 1467–1474.
[66] D. L. Boger, R. S. Ozer, C.-M. Andersson, *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1903–1908.
[67] D. L. Boger, P. Ducray, W. Chai, W. Jiang, J. Goldberg, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2339–2344.
[68] D. L. Boger, J. Goldberg, W. Jiang, W. Chai, P. Ducray, J. K. Lee, R. S. Ozer, C.-M. Andersson, *Bioorg. Med. Chem.* **1998**, *6*, 1347–1378.
[69] "Multi-step solution-phase combinatorial chemistry": D. L. Boger, J. Goldberg in *Combinatorial Chemistry: A Practical Approach* (Ed.: H. Fenniri), Oxford University Press, Oxford, UK, **2000**, pp. 303–326.
[70] D. L. Boger, B. E. Fink, M. P. Hedrick, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1019–1020.
[71] D. L. Boger, J. Goldberg, S. Satoh, Y. Ambroise, S. B. Cohen, P. K. Vogt, *Helv. Chim. Acta* **2000**, *83*, 1825–1845.
[72] D. L. Boger, J. K. Lee, *J. Org. Chem.* **2000**, *65*, 5996–6000.
[73] D. L. Boger, J. Goldberg, *Bioorg. Med. Chem.* **2001**, *9*, 557–562.
[74] D. L. Boger, H. Keim, B. Oberhauser, E. P. Schreiner, C. A. Foster, *J. Am. Chem. Soc.* **1999**, *121*, 6197–6205.
[75] D. L. Boger, Y. Chen, C. A. Foster, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1741–1744.
[76] Y. Chen, M. Bilban, C. A. Foster, D. L. Boger, *J. Am. Chem. Soc.* **2002**, *124*, 5431–5440.
[77] S. Silletti, T. Kessler, J. Goldberg, D. L. Boger, D. A. Cheresh, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 119–124.
[78] D. L. Boger, J. Goldberg, S. Silletti, T. Kessler, D. A. Cheresh, *J. Am. Chem. Soc.* **2001**, *123*, 1280–1288.
[79] Y. Ambroise, B. Yaspan, M. H. Ginsberg, D. L. Boger, *Chem. Biol.* **2002**, *9*, 1219–1226.
[80] T. Berg, S. B. Cohen, J. Desharnais, C. Sonderegger, D. J. Maslyar, J. Goldberg, D. L. Boger, P. K. Vogt, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3830–3835.
[81] D. L. Boger, M. Searcey, W. C. Tse, Q. Jin, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 495–498.
[82] J. Goldberg, Q. Jin, Y. Ambroise, S. Satoh, J. Desharnais, K. Capps, D. L. Boger, *J. Am. Chem. Soc.* **2002**, *124*, 544–555.
[83] C. R. Woods, T. Ishii, B. Wu, K. W. Bair, D. L. Boger, *J. Am. Chem. Soc.* **2002**, *124*, 2148–2152.
[84] C. R. Woods, T. Ishii, D. L. Boger, *J. Am. Chem. Soc.* **2002**, *124*, 10676–10682.
[85] C. R. Woods, N. Faucher, B. Eschgfaller, K. W. Bair, D. L. Boger, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2647–2650.
[86] D. L. Boger, M. A. Dechantsreiter, T. Ishii, B. E. Fink, M. P. Hedrick, *Bioorg. Med. Chem.* **2000**, *8*, 2049–2057.
[87] D. L. Boger, B. E. Fink, M. P. Hedrick, *J. Am. Chem. Soc.* **2000**, *122*, 6382–6394.
[88] D. L. Boger, H. W. Schmitt, B. E. Fink, M. P. Hedrick, *J. Org. Chem.* **2001**, *66*, 6654–6661.
[89] D. L. Boger, *Chemtracts: Org. Chem.* **1991**, *4*, 329–349.
[90] D. L. Boger, *Acc. Chem. Res.* **1995**, *28*, 20–29.
[91] D. L. Boger, D. S. Johnson, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3642–3649.
[92] D. L. Boger, D. S. Johnson, *Angew. Chem.* **1996**, *108*, 1542–1580; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1438–1474.
[93] D. L. Boger, R. M. Garbaccio, *Bioorg. Med. Chem.* **1997**, *5*, 263–276.
[94] D. L. Boger, R. M. Garbaccio, *Acc. Chem. Res.* **1999**, *32*, 1043–1052.
[95] D. L. Boger, H. Cai, *Angew. Chem.* **1999**, *111*, 470–500; *Angew. Chem. Int. Ed.* **1999**, *38*, 448–476.
[96] S. E. Wolkenberg, D. L. Boger, *Chem. Rev.* **2002**, *102*, 2477–2495.
[97] D. L. Boger, B. E. Fink, S. R. Brunette, W. C. Tse, M. P. Hedrick, *J. Am. Chem. Soc.* **2001**, *123*, 5878–5891.
[98] D. L. Boger, W. C. Tse, *Bioorg. Med. Chem.* **2001**, *9*, 2511–2518.
[99] D. Leung, C. Hardouin, D. L. Boger, B. F. Cravatt, *Nat. Biotechnol.* **2003**, *23*, 687–691.
[100] A. G. Cochran, *Chem. Biol.* **2000**, *7*, R85–R94.
[101] A. G. Cochran, *Curr. Opin. Chem. Biol.* **2001**, *5*, 654–659.
[102] D. L. Boger, *Bioorg. Med. Chem.* **2003**, *11*, 1607—1613.
[103] H. M. Geysen, R. H. Meloen, S. J. Barteling, *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 3998–4002.
[104] H. M. Geysen, S. J. Barteling, R. H. Meleon, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 178–182.
[105] R. A. Houghten, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 5131–5135.

[106] J. K. Scott, G. P. Smith, *Science* **1990**, *249*, 386–390.
[107] G. P. Smith, V. A. Petrenko, *Chem. Rev.* **1997**, *97*, 391–410.
[108] R. Frank, *Tetrahedron* **1992**, *48*, 9217–9232. See also: R. Frank, *Nucleic Acids Res.* **1983**, *11*, 4365.
[109] Cut and combine membrane-supported synthesis: F. Dittrich, W. Tegge, R. Frank, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2351–2356.
[110] Á. Furka, F. Sebestyén, M. Asgedom, G. Dibó, *Int. J. Peptide Prot. Res.* **1991**, *37*, 487–493; Á. Furka, F. Sebestyén, M. Asgedom, G. Dibó, *Abstr. 10th Intl. Symp. Med. Chem.*, Budapest, Ungarn, **1988**, S. 288.
[111] F. Sebestyén, G. Dibó, A. Kovács, Á. Furka, *Bioorg. Med. Chem. Lett.* **1993**, *3*, 413–418.
[112] Á. Furka, *Drug Dev. Res.* **1995**, *36*, 1–12.
[113] R. A. Houghten, C. Pinilla, S. E. Blondelle, J. R. Appel, C. T. Dooley, J. H. Cuervo, *Nature* **1991**, *354*, 84–86.
[114] K. S. Lam, S. E. Salmon, E. M. Hersh, V. J. Hruby, W. M. Kazmierski, R. J. Knapp, *Nature* **1991**, *354*, 82–84.
[115] S. E. Salmon, K. S. Lam, M. Lebl, A. Kandola, P. S. Khattri, S. Wade, M. Pátek, P. Kocis, V. Krchòák, D. Thorpe, S. Felder, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 11708–11712.
[116] K. S. Lam, M. Lebl, V. Krchòák, *Chem. Rev.* **1997**, *97*, 411–448.
[117] S. P. A. Fodor, L. J. Read, M. C. Pirrung, L. Stryer, A. T. Lu, D. Solas, *Science* **1991**, *251*, 767–773.
[118] D. J. Ecker, T. A. Vickers, R. Hanecak, V. Driver, K. Anderson, *Nucleic Acids Res.* **1993**, *21*, 1853–1856.
[119] S. M. Freier, D. A. M. Konings, J. R. Wyatt, D. J. Ecker, *J. Med. Chem.* **1995**, *38*, 344–352.
[120] E. Erb, K. D. Janda, S. Brenner, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11422–11426.
[121] S. Brenner, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5381–5383.
[122] J. Nielsen, S. Brenner, K. D. Janda, *J. Am. Chem. Soc.* **1993**, *115*, 9812–9813.
[123] J. M. Kerr, S. C. Banville, R. N. Zuckermann, *J. Am. Chem. Soc.* **1993**, *115*, 2529–2531.
[124] V. Nikolaev, A. Stierandova, V. Krchnák, B. Seligmann, K. S. Lam, S. E. Salmon, M. Lebl, *Pept. Res.* **1993**, *6*, 161–170.
[125] E. R. Felder, G. Heizmann, I. T. Matthews, H. Rink, E. Spieser, *Mol. Diversity* **1996**, *1*, 109–112.
[126] M. H. J. Ohlmeyer, R. N. Swanson, L. W. Dillard, J. C. Reader, G. Asouline, R. Kobayashi, M. Wigler, W. C. Still, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10922–10926.
[127] H. P. Nestler, P. A. Bartlett, W. C. Still, *J. Org. Chem.* **1994**, *59*, 4723–4724.
[128] W. C. Still, *Acc. Chem. Res.* **1996**, *29*, 155–163.
[129] K. C. Nicolaou, X.-Y. Xiao, Z. Parandoosh, A. Senyei, M. P. Nova, *Angew. Chem.* **1995**, *107*, 2476–2479; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2289–2291.
[130] E. J. Moran, S. Sarshar, J. F. Cargill, M. M. Shahbaz, A. Lio, A. M. M. Mjallim, R. W. Armstrong, *J. Am. Chem. Soc.* **1995**, *117*, 10787–10788.
[131] J. W. Guiles, C. L. Lanter, R. A. Rivero, *Angew. Chem.* **1998**, *110*, 967–970; *Angew. Chem. Int. Ed.* **1998**, *37*, 926–932.
[132] A. R. Vaino, K. D. Janda, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7692–7696.
[133] R. Frank, S. Hoffmann, M. Kiess, H. Lahmann, W. Tegge, C. Behn, H. Gausepohl, *Combinatorial Peptide and Nonpeptide Libraries* (Ed.: G. Jung), VCH, Weinheim, **1996**, pp. 363–386.
[134] C. M. Niemeyer, D. Blohm, *Angew. Chem.* **1999**, *111*, 3039–3043; *Angew. Chem. Int. Ed.* **1999**, *38*, 2865–2869.
[135] G. MacBeath, A. N. Koehler, S. L. Schreiber, *J. Am. Chem. Soc.* **1999**, *121*, 7967–7968.
[136] K. S. Lam, M. Renil, *Curr. Opin. Chem. Biol.* **2002**, *6*, 353–358.
[137] M. J. Plunkett, J. A. Ellman, *J. Org. Chem.* **1997**, *62*, 2885–2893.
[138] A. B. Reitz, *Curr. Opin. Drug Discovery Dev.* **1999**, *2*, 358–364.
[139] P. Blaney, R. Grigg, V. Sridharan, *Chem. Rev.* **2002**, *102*, 2607–2624.
[140] J. I. Crowley, H. Rapoport, *J. Am. Chem. Soc.* **1970**, *92*, 6363–6365.
[141] S. H. DeWitt, J. S. Kiely, C. J. Stankovic, M. C. Schroeder, D. M. R. Cody, M. R. Pavia, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6909–6913.
[142] A. R. Brown, D. C. Rees, Z. Rankovic, J. R. Morphy, *J. Am. Chem. Soc.* **1997**, *119*, 3288–3295.
[143] G. W. Kenner, J. R. McDermott, R. C. Sheppard, *J. Chem. Soc. D* **1971**, 636–637.
[144] L. M. Gayo, M. J. Sato, *Tetrahedron Lett.* **1997**, *38*, 211–214.
[145] B. J. Backes, J. A. Ellman, *J. Org. Chem.* **1999**, *64*, 2322–2330.
[146] G. T. Bourne, W. D. F. Meutermans, M. L. Smythe, *Tetrahedron Lett.* **1999**, *40*, 7271–7274.
[147] K. J. Jensen, J. Alsina, M. F. Songster, J. Vágner, F. Albericio, G. Barany, *J. Am. Chem. Soc.* **1998**, *120*, 5441–5452.
[148] T. Masquelin, D. Sprenger, R. Baer, F. Gerber, Y. Mercadal, *Helv. Chim. Acta* **1998**, *81*, 646–660.
[149] C. O. Kappe, *Bioorg. Med. Chem. Lett.* **2000**, *10*, 49–51.
[150] A. Ganesan, *Angew. Chem.* **1998**, *110*, 2989–2992; *Angew. Chem. Int. Ed.* **1998**, *37*, 2828–2831.
[151] J.-M. Lehn, *Chem. Eur. J.* **1999**, *5*, 2455–2463.
[152] B. Klekota, B. L. Miller, *Trends Biotechnol.* **1999**, *17*, 205–209.
[153] G. R. L. Cousins, S.-A. Poulsen, J. K. M. Sanders, *Curr. Opin. Chem. Biol.* **2000**, *4*, 270–279.
[154] J.-M. Lehn, A. V. Eliseev, *Science* **2001**, *291*, 2331–2332.
[155] P. W. Smith, J. Y. Q. Lai, A. R. Whittington, B. Cox, J. G. Houston, C. H. Stylli, M. N. Banks, P. R. Tiller, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 2821–2824.
[156] T. Carell, E. A. Wintner, A. Bashir-Hashemi, J. Rebek, Jr., *Angew. Chem.* **1994**, *106*, 2159–2162; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2059–2061.
[157] T. Carell, E. A. Wintner, A. J. Sutherland, J. Rebek, Jr., Y. M. Dunayevskiy, P. Vouros, *Chem. Biol.* **1995**, *2*, 171–183.
[158] Y. Dunayevskiy, P. Vouros, T. Carell, E. A. Wintner, J. Rebek, Jr., *Anal. Chem.* **1995**, *67*, 2906–2915.
[159] G. W. Shipps, Jr., U. P. Spitz, J. Rebek, Jr., *Bioorg. Med. Chem.* **1996**, *4*, 655–657.
[160] M. C. Pirrung, J. Chen, *J. Am. Chem. Soc.* **1995**, *117*, 1240–1245. M. C. Pirrung, J. H.-L. Chau, J. Chen, *Chem. Biol.* **1995**, *2*, 621–626.
[161] J. Nielsen, P. H. Rasmussen, *Tetrahedron Lett.* **1996**, *37*, 3351–3354.
[162] I. Ugi, A. Domling, W. Horl, *Endeavour* **1994**, *18*, 115–122.
[163] R. W. Armstrong, A. P. Combs, P. A. Tempest, S. D. Brown, T. A. Keating, *Acc. Chem. Res.* **1996**, *29*, 123–131.
[164] W. C. Ripka, G. Barker, J. Krakover, *Drug Discovery Today* **2001**, *6*, 471–477.
[165] "Inverse Solid Phase Synthesis": L. H. Caporale (Combichem, Inc.), US 5767238, **1998**; [*Chem. Abstr.* **1998**, *129*, 67693], filed 6/7/1995 (US application 95-483143).
[166] D. L. Boger, C. E. Brotherton, *J. Org. Chem.* **1984**, *49*, 4050–4055.
[167] D. L. Boger, M. D. Mullican, *J. Org. Chem.* **1984**, *49*, 4033–4044.
[168] D. L. Boger, M. D. Mullican, *J. Org. Chem.* **1984**, *49*, 4045–4050.
[169] S. L. Schreiber, *Science* **2000**, *287*, 1964–1969; L. Weber, *Curr. Opin. Chem. Biol.* **2000**, *4*, 295–302.
[170] B. Seed, *Chem. Biol.* **1994**, *1*, 125–129.
[171] A. Whitty, C. W. Borysenko, *Chem. Biol.* **1999**, *6*, R107–R118.
[172] Reviews: R. A. Houghten, C. Pinilla, J. R. Appel, S. Blondelle, C. T. Dooley, J. Eichler, A. Nefzi, J. M. Ostresh, *J. Med. Chem.* **1999**, *42*, 3743–3778; C. Pinella, J. R. Appel, E. Borras, R. A. Houghten, *Nat. Med.* **2003**, *9*, 118–122.

[173] C. Pinilla, J. R. Appel, P. Blanc, R. A. Houghten, *Biotechniques* **1992**, *13*, 901–905; C. T. Dooley, R. A. Houghten, *Life Sci.* **1993**, *52*, 1509–1517.

[174] Omission libraries: E. Câmpian, M. L. Peterson, H. H. Saneii, Á. Furka, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 2357–2362; T. Carell, E. A. Wintner, J. Rebek, Jr., *Angew. Chem.* **1994**, *106*, 2162–2164; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2061–2064.

[175] R. A. Houghten, C. T. Dooley, *Bioorg. Med. Chem. Lett.* **1993**, *3*, 405–412.

[176] Y. M. Dunayevskiy, J.-J. Lai, C. Quinn, F. Talley, P. Vouros, *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1178–1184. Y.-Z. Zhao, R. B. van Breemen, D. Nikolic, C.-R. Huang, C. P. Woodbury, A. Schilling, D. L. Venton, *J. Med. Chem.* **1997**, *40*, 4006–4012.

[177] P. J. Hajduk, R. P. Meadows, S. W. Fesik, Q. *Rev. Biophys.* **1999**, *32*, 211–240.

[178] M. Pellecchia, D. S. Sem, K. Wüthrich, *Nat. Rev. Drug Discovery* **2002**, *1*, 211–219.

[179] Y.-H. Chu, D. P. Kirby, B. L. Karger, *J. Am. Chem. Soc.* **1995**, *117*, 5419–5420; A. V. Eliseev, *Curr. Opin. Drug Discovery Dev.* **1998**, *1*, 106–115.

[180] M. M. Hann, A. R. Leach, G. Harper, *J. Chem. Inf. Comput. Sci.* **2001**, *41*, B56–B64.

[181] P. L. Toogood, *J. Med. Chem.* **2002**, *45*, 1543–1558. M. W. Peczuh, A. D. Hamilton, *Chem. Rev.* **2000**, *100*, 2479–2494.

[182] R. Zutshi, M. Brickner, J. Chmielewski, *Curr. Opin. Chem. Biol.* **1998**, *2*, 62–66.

[183] C.-H. Heldin, *Cell* **1995**, *80*, 213–223.

[184] J. Schlessinger, *Cell* **2002**, *110*, 669–672.

[185] N. Stahl, G. D. Yancopoulos, *Cell* **1993**, *74*, 587–590.

[186] Further reviews: H. Ogiso, R. Ishitani, O. Nureki, S. Fukai, M. Yamanaka, J.-H. Kim, K. Saito, A. Sakamoto, M. Inoue, M. Shirouzu, S. Yokoyama, *Cell* **2002**, *110*, 775–787; A. L. F. Mui, A. Miyajima, *Prog. Growth Factor Res.* **1994**, *5*, 15–35; C.-H. Heldin, A. Oestman, *Cytokine Growth Factor Rev.* **1996**, *7*, 3–10; C.-H. Heldin, *Cancer Surv.* **1996**, *27*, 7–24; M. A. Lemmon, J. Schlessinger, *Methods Mol. Biol.* **1998**, *84*, 49–71; K. Schulze-Osthoff, D. Ferrari, M. Los, S. Wesselborg, M. E. Peter, *Eur. J. Biochem.* **1998**, *254*, 439–459; M. Onishi, T. Nosaka, T. Kitamura, *Int. Rev. Immunol.* **1998**, *16*, 617–634; F. P. Ottensmeyer, D. R. Beniac, R. Z.-T. Luo, C. C. Yip, *Biochemistry* **2000**, *39*, 12103–12112; B. M. Baker, D. C. Wiley, *Immunity* **2001**, *14*, 681–692; P. T. Daniel, T. Wieder, I. Sturm, K. Schulze-Osthoff, *Leukemia* **2001**, *15*, 1022–1032; S. R. George, B. F. O'Dowd, S. P. Lee, *Nat. Rev. Drug Discovery* **2002**, *1*, 808–820; R. C. Patel, U. Kumar, D. C. Lamb, J. S. Eid, M. Rocheville, M. Grant, A. Rani, T. Hazlett, S. C. Patel, E. Gratton, Y. C. Patel, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 3294–3299.

[187] J. N. Ihle, B. A. Witthuhn, F. W. Quelle, K. Yamamoto, W. E. Thierfelder, B Kreider, O. Silvennoinen, *Trends Biochem. Sci.* **1994**, *19*, 222–227.

[188] P. Lamb, H. M. Seidel, R. B. Stein, J. Rosen, *Annu. Rep. Med. Chem.* **1996**, *31*, 269–278; A. H. Brivanlou, J. E. Darnell, Jr., *Science* **2002**, *295*, 813–818.

[189] D. S. Aaronson, C. M. Horvath, *Science* **2002**, *296*, 1653–1655.

[190] Additional reviews K. Muegge, S. K. Durum, *Cytokine* **1990**, *2*, 1–8; L. A. Neuhold, B. Wold, *Cell* **1993**, *74*, 1033–1042. E. C. B. Jørgensen, H. Autrup, *Carcinogenesis* **1996**, *17*, 435–441; D. S. Latchman, *Int. J. Biochem. Cell Biol.* **1997**, *29*, 1305–1312; T. Hoey, *Science* **1997**, *278*, 1578–1579; P. A. Clemons, *Curr. Opin. Chem. Biol.* **1999**, *3*, 112–115; A. Stephanou, D. S. Latchman, *Gene Expression* **1999**, *7*, 311–319; Z. Yu, D. Zhou, G. Cheng, M. P. Mattson, *J. Mol. Neurosci.* **2000**, *15*, 31–44; P. S. Subramaniam, B. A. Torres, H. M. Johnson, *Cytokine* **2001**, *15*, 175–187; H. van Dam, M. Castellazzi, *Oncogene* **2001**, *20*, 2453–2464; T. E. Battle, D. A. Frank, *Curr. Mol. Med.* **2002**, *2*, 381–392.

[191] Reviews: a) J. D. Klemm, S. L. Schreiber, G. R. Crabtree, *Annu. Rev. Immunol.* **1998**, *16*, 569–592; b) S. L. Schreiber, G. R. Crabtree, *Harvey Lect.* **1997**, *91*, 99–114; c) G. R. Crabtree, S. L. Schreiber, *Trends Biochem. Sci.* **1996**, *21*, 418–422; d) D. J. Austin, G. R. Crabtree, S. L. Schreiber, *Chem. Biol.* **1994**, *1*, 131–136.

[192] a) D. M. Spencer, T. J. Wandless, S. L. Schreiber, G. R. Crabtree, *Science* **1993**, *262*, 1019–1024; b) D. M. Spencer, P. J. Belshaw, L. Chen, S. N. Ho, F. Randazzo, G. R. Crabtree, S. L. Schreiber, *Curr. Biol.* **1996**, *6*, 839–847; c) P. J. Belshaw, S. N. Ho, G. R. Crabtree, S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 4604–4607; d) B. R. Stockwell, S. L. Schreiber, *Curr. Biol.* **1998**, *8*, 761–770.

[193] S. N. Ho, S. R. Biggar, D. M Spencer, S. L. Schreiber, G. R. Crabtree, *Nature* **1996**, *382*, 822–826.

[194] R. Briesewitz, G. T. Ray, T. J. Wandless, G. R. Crabtree, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 1953–1958.

[195] J. Huang, S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 13396–13401.

[196] K. Koide, J. M. Finkelstein, Z. Ball, G. L. Verdine, *J. Am. Chem. Soc.* **2001**, *123*, 398–408.

[197] M. Liuzzi, R. Deziel, N. Moss, P. Beaulieu, A. M. Bonneau, C. Bousquet, J. G. Chafouleas, M. Garneau, J. Jaramillo, R. L. Krogsrud, *Nature* **1994**, *372*, 695–698.

[198] K. McMillan, M. Adler, D. S. Auld, J. J. Baldwin, E. Blasko, L. J. Browne, D. Chelsky, D. Davey, R. Dolle, K. A. Eagen, S. Erickson, R. I. Feldman, C. B. Glaser, C. Mallari, M. M. Morrissey, M. H. Ohlmeyer, G. Pan, J. F. Parkinson, G. B. Phillips, M. A. Polokoff, N. H. Sigal, R. Vergona, M. Whitlow, T. A. Young, J. J. Devlin, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 1506–1511.

[199] a) M. D. Shultz, M. J. Bowman, Y.-W. Ham, X. Zhao, G. Tora, J. Chmielewski, *Angew. Chem.* **2000**, *112*, 2822–2825; *Angew. Chem. Int. Ed.* **2000**, *39*, 2710–2713; b) X. Fan, G. R. Flentke, D. H. Rich, *J. Am. Chem. Soc.* **1998**, *120*, 8893–8894.

[200] a) L. Zhao, M. K. O'Reilly, M. D. Schultz, J. Chmielewski, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1175–1177; b) R. G. Maroun, S. Gayet, M. S. Benleulmi, H. Porumb, L. Zararian, H. Merad, H. Leh, J.-F. Mouscadet, F. Troalen, S. Fermandjian, *Biochemistry* **2001**, *40*, 13840–13848.

[201] B. P. Orner, J. T. Ernst, A. D. Hamilton, *J. Am. Chem. Soc.* **2001**, *123*, 5382–5383.

[202] A. Degterev, A. Lugovsky, M. Cardone, B. Mulley, G. Wagner, T. Mitchison, J. Yuan, *Nat. Cell Biol.* **2001**, *3*, 173–182.

[203] S.-P. Tzung, K. M. Kim, G. Basañez, C. D. Giedt, J. Simon, J. Simmerberg, K. Y. J. Zhang, D. M. Hockenbery, *Nat. Cell Biol.* **2001**, *3*, 183–191.

[204] J.-L. Wang, D. Liu, Z.-J. Zhang, S. Shan, X. Han, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7124–7129.

[205] a) O. Kutzki, H. S. Park, J. T. Ernst, B. P. Orner, H. Yin, A. D. Hamilton, *J. Am. Chem. Soc.* **2002**, *124*, 11838–11839; b) J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem.* **2003**, *115*, 553–557; *Angew. Chem. Int. Ed.* **2003**, *42*, 535–539.

[206] a) S. J. Duncan, S. Gruschow, D. H. Williams, C. McNicholas, R. Purewal, M. Hajek, M. Gerlitz, S. Martin, S. K. Wrigley, M. Moore, *J. Am. Chem. Soc.* **2001**, *123*, 554–560; b) R. Stoll, C. Renner, S. Hansen, S. Palme, C. Klein, A. Belling, W. Zeslawski, M. Kamionka, T. Rehm, P. Muhlhahn, R. Schumacher, F. Hesse, B. Kaluza, W. Voelter, R. A. Engh, T. A. Holak, *Biochemistry* **2001**, *40*, 336–344.

[207] C. Garcia-Echeverria, P. Chene, M. J. Blommer, P. Furet, *J. Med. Chem.* **2000**, *43*, 3205–3208.

[208] J. B. Owolabi, G. Rizkalla, A. Tehim, G. M. Ross, R. J. Riopelle, R. Kamboj, M. Ossipov, D. Bian, S. Wegert, F. Porreca, D. K. H. Lee, *J. Pharmacol. Exp. Ther.* **1999**, *289*, 1271–1276.

[209] M. Baba, O. Nishimura, N. Kanzaki, M. Okamoto, H. Sawada, Y. Iizawa, M. Shiraishi, Y. Aramaki, K. Okonogi, Y. Ogawa, K. Megurp, M. Fujino, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5698–5702.

[210] a) J. W. Tilley, L. Chen, D. C. Fry, S. D. Emerson, G. D. Powers, D. Biondi, T. Varnell, R. Trilles, R. Guthrie, F. Mennona, G. Kaplan, R. A. LeMahieu, M. Carson, R.-J. Han, C.-M. Liu, R. Palermo, G. Ju, *J. Am. Chem. Soc.* **1997**, *119*, 7589–7590; b) A. C. Braisted, J. D. Oslob, W. L. Delano, J. Hyde, R. S. McDowell, N. Waal, C. Yu, M. R. Arkin, B. C. Raimundo, *J. Am. Chem. Soc.* **2003**, *125*, 3714–3715.

[211] R. Sarabu, J. P. Cooper, C. M. Cook, P. Gillespie, A. V. Perrotta, G. L. Olson, *Drug Des. Discovery* **1998**, *15*, 191–198.

[212] S. E. Cwirla, P. Balasubramanian, D. J. Duffin, C. R. Wagstrom, C. M. Gates, S. C. Singer, A. M. Davis, R. L. Tansik, L. C. Mattheakis, C. M. Boytos, P. J. Schatz, D. P. Baccanari, N. C. Wrighton, R. W. Barrett, W. J. Dower, *Science* **1997**, *276*, 1696–1699.

[213] T. Kimura, H. Kaburaki, T. Tsujino, Y. Ikeda, H. Kato, Y. Watanabe, *FEBS Lett.* **1998**, *428*, 250–254.

[214] S. S. Tian, P. Lamb, A. G. King, S. G. Miller, L. Kessler, J. I. Luengo, L. Averill, R. K. Johnson, J. G. Gleason, L. M. Pelus, S. B. Dillon, J. Rosen, *Science* **1998**, *281*, 257–259.

[215] a) A. Finch, A. K. Wong, N. J. Paczkowski, S. K. Wadi, D. J. Craik, D. P. Fairlie, S. M. Taylor, *J. Med. Chem.* **1999**, *42*, 1965–1974; b) A. J. Strachen, T. M. Woodruff, G. Haaima, D. P. Fairlie, S. M. Taylor, *J. Immunol.* **2000**, *164*, 6560–6565.

[216] J. Liu, J. D. Farmer, W. S. Lane, J. Friedman, I. Weissman, S. L. Schreiber, *Cell* **1991**, *66*, 807–815.

[217] H. M. Ke, L. D. Zydowsky, J. Liu, C. T. Walsh, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 9483–9487.

[218] a) E. J. Brown, M. W. Albers, T. B. Shin, K. Ichikawa, C. T. Keith, W. S. Lane, S. L. Schreiber, *Nature* **1994**, *369*, 756–758; b) J. Choi, J. Chen, S. L. Schreiber, J. Clardy, *Science* **1996**, *273*, 239–242.

[219] N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 905–920; N. Ban, P. Nissen, J. Hansen, P. B. Moore, T. A. Steitz, *Science* **2000**, *289*, 920–930.

[220] a) Reviews: T. Scheibel, J. Buchner, *Biochem. Pharmacol.* **1998**, *56*, 675–682; J. S. Isaacs, W. Xu, L. Neckers, *Cancer Cell*, **2003**, *3*, 213–217; b) G. Chiosis, M. N. Timaul, B. Lucas, P. M. Minster, F. F. Zheng, L. Sepp-Lorenzino, N. Rosen, *Chem. Biol.* **2001**, *8*, 289–299.

[221] a) Reviews: K. H. Downing, *Emerging Ther. Targets* **2000**, *4*, 219–237; J. R. Peterson, T. J. Mitchison, *Chem. Biol.* **2002**, *9*, 1275–1285; b) S. J. Haggarty, T. U. Mayer, D. T. Miyamoto, R. Fathi, R. W. King, T. J. Mitchison, S. L. Schreiber, *Chem. Biol.* **2000**, *7*, 275–286.

[222] Review: Y. Pommier, *Biochimie* **1998**, *80*, 255–270.

[223] Reviews: a) B. K. Hubbard, C. T. Walsh, *Angew. Chem.* **2003**, *115*, 752–789; *Angew. Chem. Int. Ed.* **2003**, *42*, 730–765; b) D. G. McCafferty, P. Cudic, M. K. Yu, D. C. Behenna, R. Kruger, *Curr. Opin. Chem. Biol.* **1999**, *3*, 672–680.

[224] a) J. A. Katzenellenbogen, B. S. Katzenellenbogen, *Chem. Biol.* **1996**, *3*, 529–536; b) D. Santiso-Mere, D. P. McDonnell, *Chem. Today*, **1994**, *12*, 29–36.

[225] a) T. K. Sawyer, *Biopolymers* **1998**, *47*, 243–261; b) J. K. Chen, S. L. Schreiber, *Bioorg. Med. Chem. Lett.* **1994**, *4*, 1755–1760; c) M. Sudol, T. Hunter, *Cell* **2000**, *103*, 1001–1004; d) A. Y. Hung, M. Sheng, *J. Biol. Chem.* **2002**, *277*, 5699–5702; M. Zhang, W. Wang, *Acc. Chem. Res.* **2003**, *36*, 530–538.

[226] J.-M. Chen, R. T. Aimes, G. R. Ward, G. L. Youngleib, J. P. Quigley, *J. Biol. Chem.* **1991**, *266*, 5113–5121.

[227] V.-M. Kahari, U. Saarialho-Kere, *Ann. Med.* **1999**, *31*, 34–45.

[228] M. A. Rudek, J. Venitz, W. D. Figg, *Pharmacotherapy* **2002**, *22*, 705–720.

[229] M. Lein, K. Jung, B. Ortel, C. Stephan, W. Rothaug, R. Juchem, M. Johannsen, S. Deger, D. Schnorr, S. Loening, H.-W. Krell, *Oncogene* **2002**, *21*, 2089–2096.

[230] L. M. Coussens, B. Fingleton, L. M. Matrisian, *Science* **2002**, *295*, 2387–2392.

[231] A. Matter, *Drug Discovery Today* **2001**, *6*, 1005–1024.

[232] For a review on integrins, see: R. O. Hynes, *Cell* **2002**, *110*, 673–687.

[233] P. C. Brooks, R. A. F. Clark, D. A. Cheresh, *Science* **1994**, *264*, 569–571.

[234] P. C. Brooks, A. M. P. Montgomery, M. Rosenfeld, R. A. Reisfeld, T. Hu, G. Klier, D. A. Cheresh, *Cell* **1994**, *79*, 1157–1164.

[235] F. Carreiras, B. Thiebot, J. Leroy-Dudal, S. Maubant, M.-F. Breton, H. Darbeida, *Int. J. Cancer* **2002**, *99*, 800–808.

[236] N. E. Tsopanoglou, P. Andriopoulou, M. E. Maragoudakis, *Am. J. Physiol.* **2002**, *283*, C1501–C1510.

[237] M. A. Arnaout, *Immunol. Rev.* **2002**, *186*, 125–140.

[238] J. Takagi, B. M. Petre, T. Walz, T. A. Springer, *Cell* **2002**, *110*, 599–611.

[239] For a review of endogenous ligand binding to integrins, see: E. F. Plow, T. A. Haas, L. Zhang, J. Loftus, J. W. Smith, *J. Biol. Chem.* **2000**, *275*, 21785–21788.

[240] E. Ruoslahti, *Annu. Rev. Cell Biol.* **1996**, *12*, 697–715.

[241] R. Haubner, R. Gratias, B. Diefenbach, S. L. Goodman, A. Jonczyk, H. Kessler, *J. Am. Chem. Soc.* **1996**, *118*, 7461–7472.

[242] R. Bruck, H. Shirin, R. Hershkoviz, O. Lider, G. Kenet, H. Aeed, Z. Matas, L. Zaidel, Z. Halpern, *Gut* **1997**, *40*, 133–138.

[243] R. Pasqualini, E. Koivunen, E. Ruoslahti, *Nat. Biotechnol.* **1997**, *15*, 542–546.

[244] A. Giannis, F. Rübsam, *Angew. Chem.* **1997**, *109*, 606–609; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 588–590.

[245] E. Ruoslahti, *Kidney Int.* **1997**, *51*, 1413–1417.

[246] K. Sunassee, R. Vile, *Curr. Biol.* **1997**, *7*, R282–R285.

[247] C. D. Eldred, B. D. Judkins, *Prog. Med. Chem.* **1999**, *39*, 29–90.

[248] C. D. Buckley, D. Pilling, N. V. Henriquez, G. Parsonage, K. Threlfall, D. Scheel-Toellner, D. L. Simmons, A. N. Akbar, J. M. Lord, M. Salmon, *Nature* **1999**, *397*, 534–539.

[249] K. Kurohane, Y. Namba, N. Oku, *Life Sci.* **2000**, *68*, 273–281.

[250] C. D. Anuradha, S. Kanno, S. Hirano, *Cell Biol. Toxicol.* **2000**, *16*, 275–283.

[251] Y. Chen, X. Xu, S. Hong, J. Chen, N. Liu, C. B. Underhill, K. Creswell, L. Zhang, *Cancer Res.* **2001**, *61*, 2434–2438.

[252] Y. Yamamoto, Y. Tsutsumi, T. Mayumi. *Curr. Drug Targets* **2002**, *3*, 123–130.

[253] V. Brower, *Nat. Biotechnol.* **1999**, *17*, 963–968.

[254] W. D. Klohs, J. M. Hamby, *Curr. Opin. Biotechnol.* **1999**, *10*, 544–549.

[255] S. R. Kerbel, *Carcinogenesis* **2000**, *21*, 505–515.

[256] R. D. Connell, J. S. Beebe, *Expert Opin. Ther. Pat.* **2001**, *11*, 1919–1945.

[257] S. Madhusudan, A. L. Harris, *Curr. Opin. Pharmacol.* **2002**, *2*, 403–414.

[258] W. W. Li, V. W. Li, D. Tsakayannis, *New Angioth.* **2002**, 547–571.

[259] P. C. Brooks, S. Strömblad, L. C. Sanders, T. L. von Schalscha, R. T. Aimes, W. G. Stetler-Stevenson, J. P. Quigley, D. A. Cheresh, *Cell* **1996**, *85*, 683–693.

[260] S. Silletti, D. A. Cheresh, *Fibrinolysis Proteolysis* **1999**, *13*, 226–238.

[261] P. C. Brooks, S. Silletti, T. L. von Schalscha, M. Friedlander, D. A. Cheresh, *Cell* **1998**, *92*, 391–400.

[262] R. O. Hynes, *Cell* **1992**, *69*, 11–25.

[263] R. L. Juliano, S. Haskill, *J. Cell Biol.* **1993**, *120*, 577–585.

[264] T. A. Springer, *Cell* **1994**, *76*, 301–314.

[265] R. R. Lobb, M. E. Hemler, *J. Clin. Invest.* **1994**, *94*, 1722–1728.

[266] M. Stewart, M. Thiel, N. Hogg, *Curr. Opin. Cell Biol.* **1995**, *7*, 690–696.
[267] Y. Shimizu, D. M. Rose, M. H. Ginsberg, *Adv. Immunol.* **1999**, *72*, 325–380.
[268] S. Liu, M. H. Ginsberg, *J. Biol. Chem.* **2000**, *275*, 22736–22742.
[269] S. Liu, S. M. Thomas, D. G. Woodside, D. M. Rose, W. B. Kiosses, M. Pfaff, M. H. Ginsberg, *Nature* **1999**, *402*, 676–681.
[270] M. E. Hemler, P. D. Kassner, B. M. C. Chan, *Cold Spring Harbor Symp. Quant. Biol.* **1992**, *57*, 213–220.
[271] K.-C. Lin, A. C. Castro, *Curr. Opin. Chem. Biol.* **1998**, *2*, 453–457.
[272] J. W. Tilley, A. Sidduri, *Drugs Future* **2001**, *26*, 985–998.
[273] H. Yusuf-Makagiansar, M. E. Anderson, T. V. Yakovleva, J. S. Murray, T. J. Siahaan, *Med. Res. Rev.* **2002**, *22*, 146–167.
[274] B. Amati, H. Land, *Curr. Opin. Genet. Dev.* **1994**, *4*, 102–108.
[275] B. Amati, S. R. Frank, D. Donjerkovic, S. Taubert, *Biochim. Biophys. Acta* **2001**, *1471*, M135–M145.
[276] L. Ratner, S. F. Josephs, F. Wong-Staal, *Annu. Rev. Microbiol.* **1985**, *39*, 419–449.
[277] C. Lavialle, N. Modjtahedi, T. Lamonerie, T. Frebourg, R. M. Landin, N. Fossar, G. Lhomond, R. Cassingena, O. Brison, *Anticancer Res.* **1989**, *9*, 1265–1279.
[278] J. K. Field, D. A. Spandidos, *Anticancer Res.* **1990**, *10*, 1–22.
[279] K. B. Marcu, S. A. Bossone, A. J. Patel, *Annu. Rev. Biochem.* **1992**, *61*, 809–860.
[280] M. Henriksson, B. Lüscher, *Adv. Cancer Res.* **1996**, *68*, 109–182.
[281] B. Calabretta, T. Skorski, *Anti-Cancer Drug Des.* **1997**, *12*, 373–381.
[282] R. Eisenman, *Genes Dev.* **2001**, *15*, 2023–2030.
[283] W. Lutz, J. Leon, M. Eilers, *Biochim. Biophys. Acta* **2002**, *1602*, 61–71.
[284] S. K. Oster, C. S. W. Ho, E. L. Soucie, L. Z. Penn, *Adv. Cancer Res.* **2002**, *84*, 81–154.
[285] A. R. Ferré-d'Amaré, G. D. Prendergast, E. B. Ziff, S. K. Burley, *Nature* **1993**, *363*, 38–45.
[286] P. Brownlie, T. A. Ceska, M. Lamers, C. Romier, G. Stier, H. Teo, D. Suck, *Structure* **1997**, *5*, 509–520.
[287] G. C. Prendergast, D. Lawe, E. B. Ziff, *Cell* **1991**, *65*, 395–407.
[288] E. M. Blackwood, R. N. Eisenman, *Science* **1991**, *251*, 1211–1217.
[289] E. M. Blackwood, B. Lüscher, R. N. Eisenman, *Genes Dev.* **1992**, *6*, 71–80.
[290] L. Kretzner, E. M. Blackwood, R. N. Eisenman, *Nature* **1992**, *359*, 426–429.
[291] G. J. Kato, W. M. F. Lee, L. Chen, C. V. Dang, *Genes Dev.* **1992**, *6*, 81–92.
[292] B. Amati, S. Dalton, M. W. Brooks, T. D. Littlewood, G. I. Evan, H. Land, *Nature* **1992**, *359*, 423–426.
[293] B. Amati, M. W. Brooks, N. Levy, T. D. Littlewood, G. I. Evan, H. Land, *Cell* **1993**, *72*, 233–245.
[294] B. Amati, T. D. Littlewood, G. I. Evan, H. Land, *EMBO J.* **1993**, *12*, 5083–5087.
[295] M. Tsuneoka, F. Nakano, H. Ohgusu, E. Mekada, *Oncogene* **1997**, *14*, 2301–2311.
[296] D. E. Ayer, L. Kretzner, R. N. Eisenman, *Cell* **1993**, *72*, 211–222.
[297] A. S. Zervos, J. Gyuris, R. Brent, *Cell* **1993**, *72*, 223–232.
[298] V. Korinek, N. Barker, P. J. Morin, D. van Wichen, R. de Weger, K. W. Kinzler, B. Vogelstein, H. Clevers, *Science* **1997**, *275*, 1784–1787.
[299] J. Behrens, *Cancer Metastasis Rev.* **1999**, *18*, 15–30.
[300] Q. Eastman, R. Grosschedl, *Curr. Opin. Cell Biol.* **1999**, *11*, 233–240.
[301] A. Novak, S. Dedhar, *Cell. Mol. Life Sci.* **1999**, *56*, 523–537.
[302] J. J. Love, X. Li, D. A. Case, K. Giese, R. Grosschedl, P. Wright, *Nature* **1995**, *376*, 791–795.
[303] M. Aoki, A. Hecht, U. Kruse, R. Kemler, P. K. Vogt, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 139–144.
[304] S. B. Krantz, *Blood* **1991**, *77*, 419–434.
[305] M. J. Koury, M. C. Bondurant, *Eur. J. Biochem.* **1992**, *210*, 649–663.
[306] H. Youssoufian, G. Longmore, D. Neumann, A. Yoshimura, H. F. Lodish, *Blood* **1993**, *81*, 2223–2236.
[307] J. E. Damen, G. Krystal, *Exp. Hematol.* **1996**, *24*, 1455–1459.
[308] P. Foa, *Acta Haematol.* **1991**, *86*, 162–168.
[309] J. W. Fisher, *Proc. Soc. Exp. Biol. Med.* **1997**, *216*, 358–369.
[310] O. Ifudu, *Nephron* **2001**, *88*, 1–5.
[311] J. Seidenfeld, M. Piper, C. Flamm, V. Hasselblad, J. O. Armitage, C. L. Bennett, M. S. Gordon, A. E. Lichtin, J. L. Wade III, S. Woolf, N. Aronson, *J. Natl. Cancer Inst.* **2001**, *93*, 1204–1214.
[312] M. Mittelman, D. Neumann, A. Peled, P. Kanter, N. Haran-Ghera, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 5181–5186.
[313] S. S. Watowich, A. Yoshimura, G. D. Longmore, D. J. Hilton, Y. Yoshimura, H. F. Lodish, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 2140–2144.
[314] J. S. Philo, K. H. Aoki, T. Arakawa, L. O. Narhi, J. Wen, *Biochemistry* **1996**, *35*, 1681–1691.
[315] D. J. Matthews, R. S. Topping, R. T. Cass, L. B. Giebel, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 9471–9476.
[316] S. J. Frank, *Endocrinology* **2002**, *143*, 2–10.
[317] M. A. Lemmon, J. Schlessinger, *Trends Biochem. Sci.* **1994**, *19*, 459–463.
[318] B. A. Witthuhn, F. W. Quelle, O. Silvennoinen, T. Yi, B. Tang, O. Miura, J. N. Ihle, *Cell* **1993**, *74*, 227–236.
[319] N. C. Wrighton, F. X. Farrell, R. Chang, A. K. Kashyap, F. P. Barbone, L. S. Mulcahy, D. L. Johnson, R. W. Barrett, L. K. Jolliffe, W. J. Dower, *Science* **1996**, *273*, 458–463.
[320] D. L. Johnson, F. X. Farrell, F. P. Barbone, F. J. McMahon, J. Tullai, D. Kroon, J. Freedy, R. A. Zivin, L. S. Mulcahy, L. K. Joliffe, *Chem. Biol.* **1997**, *4*, 939–950.
[321] O. Livnah, E. A. Stura, D. L. Johnson, S. A. Middleton, L. S. Mulcahy, N. C. Wrighton, W. J. Dower, L. K. Joliffe, I. A. Wilson, *Science* **1996**, *273*, 464–471.
[322] I. A. Wilson, L. K. Joliffe, *Curr. Opin. Struct. Biol.* **1999**, *9*, 696–702
[323] A. M. de Vos, M. Ultsch, A. A. Kossiakoff, *Science* **1992**, *255*, 306–312.
[324] W. Somers, M. Ultsch, A. M. de Vos, A. A. Kossiakoff, *Nature* **1994**, *372*, 478–481.
[325] B. C. Cunningham, M. Ultsch, A. M. de Vos, M. G. Mulkerrin, K. R. Clauser, J. A. Wells, *Science* **1991**, *254*, 821–825.
[326] G. Fuh, B. C. Cunningham, R. Fukunaga, S. Nagata, D. V. Goeddel, J. A. Wells, *Science* **1992**, *256*, 1677–1680.
[327] B. C. Cunningham, J. A. Wells, *J. Mol. Biol.* **1993**, *234*, 554–563.
[328] J. A. Wells, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1–6.
[329] N. J. Skelton, S. Russell, F. de Sauvage, A. G. Cochran, *J. Mol. Biol.* **2002**, *316*, 1111–1125.
[330] S. A. Qureshi, R. M. Kim, Z. Konteatis, D. E. Biazzo, H. Motamedi, R. Rodrigues, J. A. Boice, J. R. Calaycay, M. A. Bednarek, P. Griffin, Y.-D. Gao, K. Chapman, D. F. Mark, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 12156–12161.
[331] O. Livnah, D. L. Johnson, E. A. Stura, F. X. Farrell, F. P. Barbone, Y. You, K. D. Liu, M. A. Goldsmith, W. He, C. D. Krause, S. Pestka, L. K. Jolliffe, I. A. Wilson, *Nat. Struct. Biol.* **1998**, *5*, 993–1004.
[332] O. Livnah, E. A. Stura, S. A. Middleton, D. L. Johnson, L. K. Jolliffe, I. A. Wilson, *Science* **1999**, *283*, 987–990.
[333] R. S. Syed, S. W. Reid, C. Li, J. C. Cheetham, K. H. Aoki, B. Liu, H. Zhan, T. D. Osslund, A. J. Chirino, J. Zhang, J. Finer-Moore, S. Elliott, K. Sitney, B. A. Katz, D. J. Matthews, J. J. Wendoloski, J. Egrie, R. M. Stroud, *Nature* **1998**, *395*, 511–516.

[334] I. Remy, I. A. Wilson, S. W. Michnick, *Science* **1999**, *283*, 990–993.
[335] Y. Choo, I. Sánchez-García, A. Klug, *Nature* **1994**, *372*, 642–645.
[336] S. Neidle, D. E. Thurston in *New Targets for Cancer Chemotherapy* (Eds.: D. J. Kerr, P. Workman), CRC Press, Boca Raton, FL, **1994**.
[337] S. Neidle, *Anti-Cancer Drug Des.* **1997**, *12*, 433–442.
[338] M. Mrksich, M. E. Parks, P. B. Dervan, *J. Am. Chem. Soc.* **1994**, *116*, 7983–7988.
[339] J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1996**, *382*, 559–561.
[340] D. Mercola, J. S. Cohen, *Cancer Gene Ther.* **1995**, *2*, 47–59.
[341] M. J. Browne, P. L. Thurlbey, *Genomes, Molecular Biology and Drug Discovery*, Academic Press, London, **1996**.
[342] H. Knudsen, P. E. Nielsen, *Nucleic Acids Res.* **1996**, *24*, 494–500.
[343] L. Good, P. E. Nielsen, *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 431–437.
[344] A. J. Gutierrez, M. D. Matteucci, D. Grant, A. Matsumura, R. W. Wagner, B. C. Froehler, *Biochemistry* **1997**, *36*, 743–748.
[345] G. W. Shipps, Jr., K. E. Pryor, J. Xian, D. A. Skyler, E. H. Davidson, J. Rebek, Jr., *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11833–11838.
[346] B. Klekota, M. H. Hammond, B. L. Miller, *Tetrahedron Lett.* **1997**, *38*, 8639–8642.
[347] B. Klekoka, B. L. Miller, *Tetrahedron* **1999**, *55*, 11687–11697.
[348] G. Werstuck, M. R. Green, *Science* **1998**, *282*, 296–298.
[349] T. Lescrinier, C. Hendrix, L. Kerremans, J. Rozenski, A. Link, B. Samyn, A. Van Aerschot, E. Lescrinier, R. Eritja, J. Van Beeumen, P. Herdewijn, *Chem. Eur. J.* **1998**, *4*, 425–433.
[350] C. Bailly, J. B. Chaires, *Bioconjugate Chem.* **1998**, *9*, 513–538.
[351] S.-Y. Chiang, J. C. Azizkhan, T. A. Beerman, *Biochemistry* **1998**, *37*, 3109–3115.
[352] D. E. Thurston, *Br. J. Cancer* **1999**, *80*, 65–85 (Suppl. 1).
[353] A. Simoncsits, M.-L. Tjornhammar, S. Wang, S. Pongor, *Genetica* **1999**, *106*, 85–92.
[354] C. M. Drain, X. Gong, V. Ruta, C. E. Soll, P. F. Chicoineau, *J. Comb. Chem.* **1999**, *1*, 286–290.
[355] V. M. Guelev, M. T. Harting, R. S. Lokey, B. L. Iverson, *Chem. Biol.* **2000**, *7*, 1–8.
[356] Md. R. Alam, M. Maeda, S. Sasaki, *Bioorg. Med. Chem.* **2000**, *8*, 465–473.
[357] Wechselwirkungen zwischen DNA und Wirkstoffen: *Methods in Molecular Biology, Vol. 90* (Ed.: K. R. Fox), Humana Press, Totowa, NJ, **1997**.
[358] H. R. Drew, A. A. Travers, *Cell* **1984**, *37*, 491–502.
[359] P. Hardenbol, J. C. Wang, M. W. Van Dyke, *Bioconjugate Chem.* **1997**, *8*, 616–620.
[360] J.-B. LePecq, C. Paoletti, *J. Mol. Biol.* **1967**, *27*, 87–106.
[361] B. S. P. Reddy, S. M. Sondhi, J. W. Lown, *Pharmacol. Ther.* **1999**, *84*, 1–111.
[362] S. Neidle, *Nat. Prod. Rep.* **2001**, *18*, 291–309.
[363] P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
[364] J. W. Trauger, E. E. Baird, M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6160–6166.
[365] R. P. Lamamie de Clairac, C. J. Seel, B. H. Geierstanger, M. Mrksich, E. E. Baird, P. B. Dervan, D. E. Wemmer, *J. Am. Chem. Soc.* **1999**, *121*, 2956–2964.
[366] J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1997**, *382*, 559–561.
[367] L. A. Dickinson, R. J. Gulizia, J. W. Trauger, E. E. Baird, D. E. Mosier, J. M. Gottesfeld, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12890–12895.
[368] Reviews: a) L. H. Caporale, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 75–82; b) R. P. Borris, *J. Ethnopharmacol.* **1996**, *51*, 29–38; c) D. G. I. Kingston in *The Practice of Medicinal Chemistry* (Ed.: C. G. Wermuth), Academic Press, London **1996**, pp. 101–116; d) G. M. Cragg, D. J. Newman, K. M. Snader, *J. Nat. Prod.* **1997**, *60*, 52–60; e) D. J. Newman, G. M. Cragg, K. M. Snader, *Nat. Prod. Rep.* **2000**, *17*, 215–234; f) Y.-Z. Shu, *J. Nat. Prod.* **1998**, *61*, 1053–1071; g) C. A. Montanari, V. da S. Bolzani, *Quim. Nova* **2001**, *24*, 105–111; h) A. B. Da Rocha, R. M. Lopes, G. Schwartsmann, *Curr. Opin. Pharmacol.* **2001**, *1*, 364–369; i) J. Mann, *Nat. Rev. Cancer* **2002**, *2*, 143–148.
[369] P. Arya, R. Joseph, D. T. H. Chou, *Chem. Biol.* **2002**, *9*, 145–156.
[370] a) J. Green, *J. Org. Chem.* **1995**, *60*, 4287–4290; b) R. Devraj, M. Cushman, *J. Org. Chem.* **1996**, *61*, 9368–9373.
[371] J. Nielsen, L. O. Lyngso, *Tetrahedron Lett.* **1996**, *37*, 8439–8442.
[372] a) T. C. Norman, N. S. Gray, J. T. Koh, P. G. Schultz, *J. Am. Chem. Soc.* **1996**, *118*, 7430–7431; b) N. S. Gray, L. Wodicka, A.-M. W. H. Thunnissen, T. C. Norman, S. Kwon, F. Herman Espinoza, D. O. Morgan, G. Barnes, S. LeClerc, L. Meijer, S.-H. Kim, D. J. Lockhart, P. G. Schultz, *Science* **1998**, *281*, 533–538; c) J. I. Armstrong, A. R. Portley, Y.-T. Chang, D. M. Nierengarten, B. N. Cook, K. G. Bowman, A. Bishop, N. S. Gray, K. M. Shokat, P. G. Schultz, C. R. Bertozzi, *Angew. Chem.* **2000**, *112*, 1359–1362; *Angew. Chem. Int. Ed.* **2000**, *39*, 1303–1306; d) Y.-T. Chang, N. S. Gray, G. R. Rosania, T. C. Kwon, R. Sarohira, M. Leost, L. Meijer, P. G. Schultz, *Chem. Biol.* **1999**, *6*, 361–375; e) G. R. Rosania, Y.-T. Chang, O. Perez, D. Sutherlin, H. Dong, D. J. Lockhart, P. G. Schultz, *Nat. Biotechnol.* **2000**, *18*, 304–308.
[373] A. Atubega, D. Maclean, C. Nguyen, E. M. Gordon, J. W. Jacobs, *Bioorg. Med. Chem.* **1996**, *4*, 1097–1106.
[374] a) K. C. Nicolaou, N. Winssinger, J. Pastor, S. Ninkovic, F. Sarabia, Y. He, D. Vourloumis, Z. Yang, T. Li, P. Giannakakou, E. Hamel, *Nature* **1997**, *387*, 268–272; b) K. C. Nicolaou, D. Vourloumis, T. Li, J. Pastor, N. Winssinger, Y. He, S. Ninkovic, F. Sarabia, H. Vallberg, F. Roschangar, N. P. King, M. R. V. Finlay, P. Giannakakou, P. Verdier-Pinard, E. Hamel, *Angew. Chem.* **1997**, *109*, 2181–2187; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2097–2103.
[375] a) S. Chen, K. D. Janda, *J. Am. Chem. Soc.* **1997**, *119*, 8724–8725; b) S. Chen, K. D. Janda, *Tetrahedron Lett.* **1998**, *39*, 3943–3946; c) K. J. Lee, A. Angulo, P. Ghazal, K. D. Janda, *Org. Lett.* **1999**, *1*, 1859–1862; d) J. A. López-Pelegrín, K. D. Janda, *Chem. Eur. J.* **2000**, *6*, 1917–1922.
[376] X. Y. Xiao, Z. Parandoosh, M. P. Nova, *J. Org. Chem.* **1997**, *62*, 6029–6033.
[377] T. R. Marsh, M. Bradley, S. J. Teague, *J. Org. Chem.* **1997**, *62*, 6199–6203.
[378] K. C. Nicolaou, N. Winssinger, J. Pastor, F. Murphy, *Angew. Chem.* **1998**, *110*, 2677–2680; *Angew. Chem. Int. Ed.* **1998**, *37*, 2534–2537.
[379] M. F. Gordeev, G. W. Luehr, H. C. Hui, E. M. Gordon, D. V. Patel, *Tetrahedron* **1998**, *54*, 15879–15890.
[380] M. Marder, H. Viola, J. A. Bacigaluppo, M. I. Colombo, C. Wasowski, C. Wolfman, J. H. Medina, E. A. Ruveda, A. C. Paladini, *Biochem. Biophys. Res. Commun.* **1998**, *249*, 481–485.
[381] K. C. Nicolaou, J. Pastor, N. Winssinger, F. Murphy, *J. Am. Chem. Soc.* **1998**, *120*, 5132–5133.
[382] a) K. C. Nicolaou, N. Winssinger, D. Vourloumis, T. Ohshima, S. Kim, J. Pfefferkorn, J.-Y. Xu, T. Li, *J. Am. Chem. Soc.* **1998**, *120*, 10814–10826; b) K. C. Nicolaou, S. Kim, J. A. Pfefferkorn, J. Xu, T. Ohshima, S. Hosokawa, D. Vourloumis, T. Li, *Angew. Chem.* **1998**, *110*, 1484–1487; *Angew. Chem. Int. Ed.* **1998**, *37*, 1418–1421.
[383] L. Bhat, Y. Liu, S. F. Victory, R. H. Himes, G. I. Georg, *Bioorg. Med. Chem. Lett.* **1998**, *8*, 3181–3186.
[384] a) D. R. Dragoli, L. A. Thompson, J. O'Brien, J. A. Ellman, *J. Comb. Chem.* **1999**, *1*, 534–539; b) L. A. Thompson, F. L.

Moore, Y.-C. Moon, J. A. Ellman, *J. Org. Chem.* **1998**, *63*, 2066–2067.
[385] A. V. Loevezijn, J. H. V. Maarseveen, K. Stegman, G. M. Visser, G.-J. Koomen, *Tetrahedron Lett.* **1998**, *39*, 4737–4740.
[386] B. Meseguer, D. Alonso-Díaz, N. Griebenow, T. Herget, H. Waldmann, *Angew. Chem.* **1999**, *111*, 3083–3087; *Angew. Chem. Int. Ed.* **1999**, *38*, 2902–2906.
[387] H. Doi, I. Hijikuro, T. Takahashi, *J. Am. Chem. Soc.* **1999**, *121*, 6749–6750.
[388] R. A. Fecik, K. E. Frank, E. J. Gentry, L. A. Mitscher, M. Shibata, *Pure Appl. Chem.* **1999**, *71*, 559–564.
[389] R. Xu, G. Greiveldinger, L. E. Marenus, A. Cooper, J. A. Ellman, *J. Am. Chem. Soc.* **1999**, *121*, 4898–4899.
[390] C. W. Lindsey, L. K. Chan, B. C. Goess, R. Joseph, M. D. Shair, *J. Am. Chem. Soc.* **2000**, *122*, 422–423.
[391] H. Wang, A. Ganesan, *J. Comb. Chem.* **2000**, *2*, 186–194.
[392] a) K. C. Nicolaou, J. A. Pfefferkorn, A. J. Roecker, G.-Q. Cao, S. Barluenga, H. J. Mitchell, *J. Am. Chem. Soc.* **2000**, *122*, 9939–9953; b) K. C. Nicolaou, J. A. Pfefferkorn, H. J. Mitchell, A. J. Roecker, S. Barluenga, G.-Q. Cao, R. I. Affleck, J. E. Lillig, *J. Am. Chem. Soc.* **2000**, *122*, 9954–9967; c) K. C. Nicolaou, J. A. Pfefferkorn, S. Barluenga, H. J. Mitchell, A. J. Roecker, G.-Q. Cao, *J. Am. Chem. Soc.* **2000**, *122*, 9968–9976.
[393] R. Maltais, M. R. Tremblay, D. Poirier, *J. Comb. Chem.* **2000**, *2*, 604–614.
[394] P. Wipf, J. T. Reeves, R. Balachandran, K. A. Giuliano, E. Hamel, B. W. Day, *J. Am. Chem. Soc.* **2000**, *122*, 9391–9395.
[395] M. B. Andrus, T. M. Turner, Z. E. Sauna, S. V. Ambudkar, *J. Org. Chem.* **2000**, *65*, 4973–4983.
[396] a) O. de Frutos, D. P. Curran, *J. Comb. Chem.* **2000**, *2*, 639–649; b) W. Zhang, Z. Luo, C. H. Chen, D. P. Curran, *J. Am. Chem. Soc.* **2002**, *124*, 10443–10450.
[397] H. E. Pelish, N. J. Westwood, Y. Feng, T. Kirchhausen, M. D. Shair, *J. Am. Chem. Soc.* **2001**, *123*, 6740–6741.
[398] R. A. Davis, A. R. Carroll, R. J. Quinn, *Aust. J. Chem.* **2001**, *54*, 355–359.
[399] I. Hijikuro, T. Doi, T. Takahashi, *J. Am. Chem. Soc.* **2001**, *123*, 3716–3722.
[400] T. Hanazawa, T. Wada, T. Masuda, S. Okamoto, F. Sato, *Org. Lett.* **2001**, *3*, 3975–3977.
[401] K. C. Nicolaou, R. Hughes, J. A. Pfefferkorn, S. Barluena, A. J. Roecker, *Chem. Eur. J.* **2001**, *7*, 4280–4295.
[402] C. Herman, C. Giammasi, A. Geyer, M. E. Maier, *Tetrahedron* **2001**, *57*, 8999–9010.
[403] a) P. Stahl, L. Kissau, R. Mazitschek, A. Huwe, P. Furet, A. Giannis, H. Waldmann, *J. Am. Chem. Soc.* **2001**, *123*, 11586–11593; b) P. Stahl, L. Kissau, R. Mazitschek, A. Giannis, H. Waldmann, *Angew. Chem.* **2002**, *114*, 1222–1226; *Angew. Chem. Int. Ed.* **2002**, *41*, 1174–1178.
[404] R. Maltais, V. Luu-The, D. Poirier, *J. Med. Chem.* **2002**, *45*, 640–653.
[405] Y.-T. Chang, J. Choi, S. Ding, E. E. Prieschl, T. Baumruker, J.-M. Lee, S.-K. Chung, P. G. Schultz, *J. Am. Chem. Soc.* **2002**, *124*, 1856–1857.
[406] D. Brohm, S. Metzger, A. Bhargava, O. Muller, F. Lieb, H. Waldmann, *Angew. Chem.* **2002**, *114*, 319–323; *Angew. Chem. Int. Ed.* **2002**, *41*, 307–311.
[407] Reviews: a) D. G. I. Kingston, D. J. Newman, *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 304–316; b) J. Nielsen, *Curr. Opin. Chem. Biol.* **2002**, *6*, 297–305; c) S. P. Manly, R. Padmanabha, S. E. Lowe, *Methods Mol. Biol.* **2002**, *190*, 153–168; d) T. Paululat, Y.-Q. Tang, S. Grabley, R. Thiericke, *Chim. Oggi* **1999**, *17*, 52–56; e) S. Bertels, S. Frormann, G. Jas, K. U. Bindseil, *Drug Discovery from Nature*, Springer, Berlin, **1999**, pp. 72–105; f) L. A. Wessjohann, *Curr. Opin. Chem. Biol.* **2000**, *4*, 303–309; g) D. G. Hall, S. Manku, F. Wang, *J. Comb. Chem.* **2001**, *3*, 125–150; h) K. C. Nicolaou, J. A. Pfefferkorn, *Biopolymers* **2001**, *60*, 171–193; i) U. Abel, C. Koch, M. Speitling, F. G. Hannsske, *Curr. Opin. Chem. Biol.* **2002**, *6*, 453–458.
[408] a) C. A. Foster, M. Dreyfuss, B. Mandak, J. G. Meingassner, H. U. Naegeli, A. Nussbaumer, L. Oberer, G. Scheel, E.-M. Swoboda, *J. Dermatol.* **1994**, *21*, 847–854; b) U. Hommel, H.-P. Weber, L. Oberer, H. U. Naegeli, B. Oberhauser, C. A. Foster, *FEBS Lett.* **1996**, *379*, 69–73.
[409] H. Itazaki, T. Fujiwara, A. Sato, Y. Kawamura, K. Matsumoto, JP 07109299, **1995**.
[410] a) D. Thanos, T. Maniatis, *Cell* **1995**, *80*, 529–532; b) T. Collins, M. A. Read, A. S. Neish, M. Z. Whitley, D. Thanos, T. Maniatis, *FASEB J.* **1995**, *9*, 899–909; c) M. J. May, S. Ghosh, *Immunol. Today* **1998**, *19*, 80–88.
[411] a) J. M. Glover, J. M. Leeds, T. G. K. Mant, D. Amin, D. L. Kisner, J. E. Zuckerman, R. S. Geary, A. A. Levin, W. R. Shanahan, Jr., *J. Pharmacol. Exp. Ther.* **1997**, *282*, 1173–1180; b) B. R. Yacyshyn, M. B. Bowen-Yacyshyn, L. Jewell, J. A. Tami, C. F. Bennett, D. L. Kisner, W. R. Shanahan, Jr., *Gastroenterology* **1998**, *114*, 1133–1142.
[412] A. F. Kavanaugh, H. Schulze-Koops, L. S. Davis, P. E. Lipsky, *Arthritis Rheum.* **1997**, *40*, 849–853.
[413] a) D. H. Boschelli, J. B. Kramer, S. S. Khatana, R. J. Sorenson, D. T. Connor, M. A. Ferin, C. D. Wright, M. E. Lesch, K. Imre, G. C. Okonkwo, D. J. Schrier, M. C. Conroy, E. Ferguson, J. Woelle, U. Saxena, *J. Med. Chem.* **1995**, *38*, 4597–4614; b) D. H. Boschelli, D. T. Connor, M. E. Lesch, D. J. Schrier, *Bioorg. Med. Chem.* **1996**, *4*, 557–562.
[414] R. W. Sullivan, C. G. Bigam, P. E. Erdman, M. S. S. Palanki, D. W. Anderson, M. E. Goldman, L. J. Ransone, M. J. Suto, *J. Med. Chem.* **1998**, *41*, 413–419.
[415] A. O. Stewart, P. A. Bhatia, C. M. McCarty, M. V. Patel, M. A. Staeger, D. L. Arendsen, I. W. Gunawardana, L. M. Melcher, G.-D. Zhu, S. A. Boyd, D. G. Fry, B. L. Cool, L. Kifle, K. Lartey, K. C. Marsh, A. J. Kempf-Grote, P. Kilgannon, W. Wisdom, J. Meyer, W. M. Gallatin, G. F. Okasinski, *J. Med. Chem.* **2001**, *44*, 988–1002.
[416] E. Di Cera, *Adv. Protein Chem.* **1998**, *51*, 59–119, and references therein.
[417] D. D. Ryu, D. H. Nam, *Biotechnol. Prog.* **2000**, *16*, 2–16.
[418] H. Lin, V. W. Cornish, *Angew. Chem.* **2002**, *114*, 4580–4606; *Angew. Chem. Int. Ed.* **2002**, *41*, 4402–4425.

JOC Article

Synthesis of Isochrysohermidin–Distamycin Hybrids

Bryan K. S. Yeung and Dale L. Boger*

Department of Chemistry and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

boger@scripps.edu

Received March 11, 2003

The synthesis of the alkylating subunit of the DNA cross-linking agent, isochrysohermidin (**2**), and its subsequent incorporation into conjugates with distamycin A (**1**) are described. The DNA binding properties of these agents were compared to that of distamycin A, using a fluorescence intercalator displacement (FID) assay.

Introduction

Gene expression is regulated by a host of inhibitor and enhancer proteins that selectively bind to specific sequences of DNA. The selective disruption of this process by small molecules, which bind to DNA in a sequence-specific manner, may provide access to new therapeutics. Among this class of agents, distamycin A is one of the most widely studied. Distamycin A, originally isolated from *Streptomyces sp.*,[1] is a minor groove binding agent with sequence specificity toward A–T rich sites within duplex DNA. Its sequence specificity and high affinity is derived from a combination of interactions including hydrogen bonding, van der Waals contacts, and electrostatic interactions of the cationic amidine side chain with the phosphate backbone of DNA.[2] The more recent discovery of 2:1 complexes,[3] their elaboration into side-by-side antiparallel γ-hairpin polyamides, and the advent of Dervan's pairing rules with the template modifications to selectively recognize G (Im vs Py) or A (Hp vs Py) provides a powerful paradigm on which to design sequence-selective DNA binding agents.[4] A number of studies have examined the consequences of incorporating nonselective alkylating agents into the distamycin structure (e.g. α-haloacylamides, nitrogen mustards).[5] A more limited series of studies have examined conjugates with selective alkylating agents (e.g. CBI,[6] duocarmycin A[7]) that combine the noncovalent binding selectivity inherent in the distamycin conjugate with the alkylation selectivity to further enhance binding selectivity and affinity.

In recent studies, we described the total synthesis of isochrysohermidin (**2**) and disclosed the first report of its interstrand DNA cross-linking properties.[8] Isolated from *Mercurialis perennis* L., both *d,l*- and *meso*-forms were found to occur naturally with the *d,l*-diastereomer unambiguously identified by X-ray crystallography.[9] The dimeric *N*-methylcarbinolamides undergo a slow ring-opening event during the interconversion of *d,l*- and *meso*-**2**. This ring-opening reaction exposes an electrophilic carbonyl capable of trapping nucleophiles within the minor groove of duplex DNA. The only nucleophile readily accessible to minor groove bound isochrysohermidin is believed to be the C2 amine of guanine. By incorporation of a single carbinolamide subunit of isochrysohermidin into distamycin, we sought to establish whether it may be possible to direct a reversible (vs irreversible)[5–7] guanine alkylation near adjacent A–T rich sites within duplex DNA (Figure 1).

Results and Discussion

The distamycin analogues were prepared by solution-phase synthesis requiring only acid/base liquid–liquid extraction protocols for the isolation and purification of the distamycin subunits. The amidine side chain found in the natural product was replaced with a *N,N*-dimethylaminopropylamine side chain to facilitate the ease of synthesis.[10] This substitution is well-documented and

(1) Arcamone, F.; Penco, S.; Orezzi, P. G.; Nicolella, V.; Pirelli, A. *Nature* **1964**, *203*, 1064.
(2) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376.
(3) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393.
(4) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
(5) For recent examples, see α-Haloacylamides: (a) Cozzi, P.; Beria, I.; Caldarelli, M.; Capolongo, L.; Geroni, C.; Mongelli, N. *Bioorg. Med. Chem. Lett.* **2000**, 10, 1269. (b) Cozzi, P. *Farmaco* **2001**, *56*, 57. (c) Cozzi, P.; Mongelli, N. *Curr. Pharm. Des.* **1998**, *4*, 181. (d) Krowicki, K.; Balzarini, J.; De Clerq, E.; Newman, R. A.; Lown, J. W. *J. Med. Chem.* **1988**, *31*, 341. (e) Baker, B. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 8266. Nitrogen mustards: (f) Wang, Y.; Wright, S. C.; Larrick, J. W. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 459. (g) Baraldi, P. G.; Romagnoli, R.; Guadix, A. E.; Pineda de la Infantas, M. J.; Gallo, M. A.; Espinosa, A.; Martinez, A.; Bingham, J. P.; Hartley, J. A. *J. Med. Chem.* **2002**, *45*, 3630. (h) Cozzi, P. C.; Beria, I.; Caldarelli, M.; Capolongo, L.; Geroni, C.; Mazzini, S.; Ragg, E. *Bioorg. Med. Chem. Lett.* **2000**, *10*, 1653. (i) Xie, G.; Gupta, R.; Lown, J. W. *Anti-Cancer Drug Des.* **1995**, *10*, 389. (j) Sigurdsson, S. T.; Rink, S. M.; Hopkins, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 12633. (k) Sondhi, S. M.; Praveen-Reddy, B. S.; Lown, J. W. *Curr. Med. Chem.* **1997**, *4*, 313.
(6) (a) Chang, A. Y.; Dervan, P. B. *J. Am. Chem. Soc.* **2000**, *122*, 4856. (b) Kumar R.; Lown, J. W. *Org. Lett.* **2002**, *4*, 1851. (c) Jia, G. F.; Iida, H.; Lown, J. W. *Heterocycl. Commun.* **1998**, *4*, 557. (d) Jia, G.; Iida, H.; Lown, J. W. *Chem. Commun.* **1999**, 119. (e) Jia, G.; Iida, H.; Lown, J. W. *Synlett* **2000**, 603. (f) Boger, D. L.; Schmitt, H.; Fink, B. E.; Hedrick, M. P. *J. Org. Chem.* **2001**, *66*, 6654.
(7) Tao, Z.-H.; Fujiwara, T.; Saito, I.; Sugiyama, H. *Angew. Chem., Int. Ed.* **1999**, *38*, 650.
(8) Boger, D. L.; Baldino, C. M. *J. Am. Chem. Soc.* **1993**, *115*, 11418.
(9) Matsui, Y.; Kawabe, C.; Matsumoto, K.; Abe, K.; Miwa, T. *Phytochemistry* **1989**, *25*, 1470.
(10) Boger, D. L.; Fink, B. E.; Hedrick, M. P. *J. Am. Chem. Soc.* **2000**, *122*, 6382.

10.1021/jo034326c CCC: $25.00
Published on Web 05/21/2003

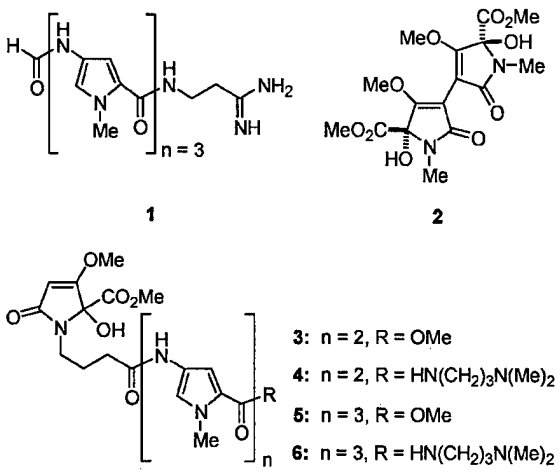
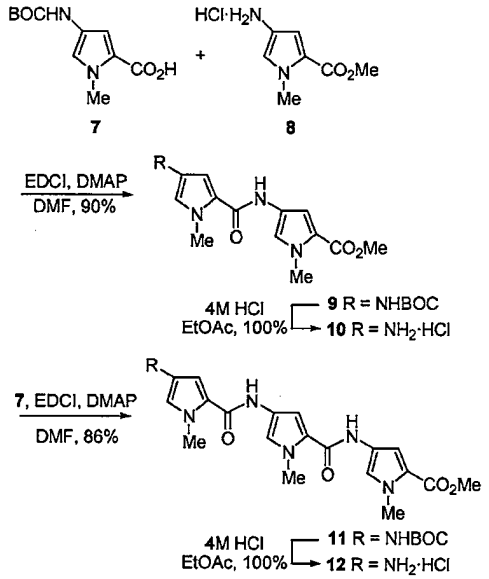


FIGURE 1. Distamycin A (**1**), isochrysohermidin (**2**), and hybrid agents (**3**–**6**).

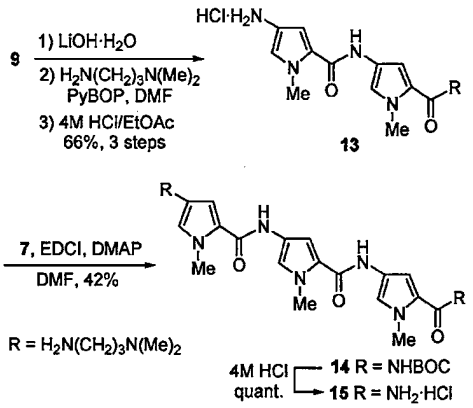
SCHEME 1



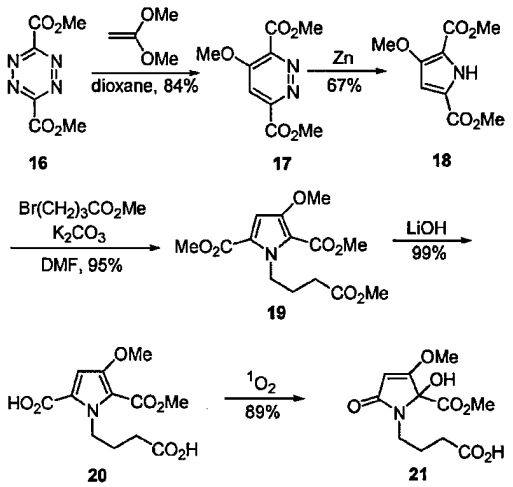
does not adversely affect minor groove binding affinity or selectivity. The distamycin subunits **10**, **12**, **13**, and **15** were prepared as previously detailed[10] by coupling **7** with **8** in the presence of EDCI and DMAP to afford **9** in 90% yield (Scheme 1). Subsequent treatment of **9** with anhydrous 4 M HCl/EtOAc followed by coupling with **7**, in the presence of EDCI/DMAP, provided the distamycin core **11** in 86% yield as detailed in our prior efforts. Subsequent treatment of peptides **9** and **11** with 4 M HCl/EtOAc removed the BOC group and provided the corresponding HCl salt of the methyl ester derivatives of the distamycin subunits (**10** and **12**).

Distamycin subunits **13** and **15**, incorporating the *N,N*-dimethylaminopropylamine side chain, were accessed from **9** as previously described.[10] Saponification with LiOH was followed by the addition of *N,N*-dimethylaminopropylamine and PyBOP to afford the corresponding adduct. Subsequent treatment with 4 M HCl/EtOAc provided **13** in 66% yield over three steps (Scheme 2). The tripeptide **14** was prepared by coupling dipeptide **13**

SCHEME 2



SCHEME 3



with **7**, in the presence of EDCI and DMAP, and provided the corresponding tripeptide 42% yield. Treatment of **14** with 4 M HCl/EtOAc removed the BOC group and provided the corresponding HCl salt of the distamycin subunit **15** in quantitative yield.

The trisubstituted pyrrole precursor to the isochrysohermidin alkylation subunit was prepared utilizing a 1,2,4,5-tetrazine → 1,2-diazine → pyrrole Diels–Alder strategy (Scheme 3).[11] An inverse electron demand Diels–Alder reaction of 1,2,4,5-tetrazine **16**[12] and dimethoxyketene acetal[13] provided the corresponding 1,2-diazine **17**[14] in 84% yield. A subsequent reductive ring contraction of the 1,2-diazine was achieved by using freshly activated zinc dust in glacial acetic acid to provide the trisubstituted pyrrole **18**[14] in 67% yield. The four-carbon tether was installed by *N*-alkylation of **18** with methyl

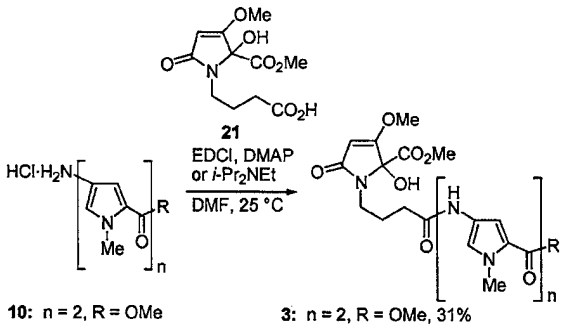
(11) For reviews: (a) Boger, D. L. *Tetrahedron* **1983**, *39*, 2869. (b) Boger, D. L. *Chem. Rev.* **1986**, *86*, 781. (c) Boger, D. L. *Bull. Soc. Chim., Belg.* **1990**, *99*, 599. (d) Boger, D. L. *Chemtracts: Org. Chem.* **1996**, *9*, 149. (e) Boger, D. L.; Weinreb, S. M. *Hetero Diels–Alder Methodology in Organic Synthesis*; Academic: San Diego, CA, 1987.

(12) Boger, D. L.; Coleman, R. S.; Panek, J. S.; Huber, F. X.; Sauer, J. *J. Org. Chem.* **1985**, *50*, 5377.

(13) (a) Corey, E. J.; Bass, J. D.; LeMahieu, R.; Mitra, R. B. *J. Am. Chem. Soc.* **1964**, *86*, 5570. (b) Zheng, Q. H.; Su, J. *Synth. Commun.* **1999**, *29*, 3467.

(14) (a) Boger, D. L.; Coleman, R. S.; Panek, J. S.; Yohannes, D. *J. Org. Chem.* **1984**, *49*, 4405. (b) Boger, D. L.; Patel, M. *J. Org. Chem.* **1988**, *53*, 1405. (c) Boger, D. L.; Baldino, C. M. *J. Org. Chem.* **1991**, *56*, 6942.

SCHEME 4



4-bromobutyrate (K_2CO_3, DMF) to furnish **19** in 95% yield. Highly selective saponification of the two sterically and electronically more accessible methyl esters provided the diacid **20** (99%). Initially, more elaborate, selective protections were anticipated to be necessary to cleanly provide **20** (e.g. use of benzyl 4-bromobutyrate). However, such selective protections/deprotections proved unnecessary and simple treatment of **19** with 2 equiv of LiOH at 25 °C provided **20** in superb conversions (99%). In the final step, a [4+2] cycloaddition of $^1O_2$ across the pyrrole followed by a low-temperature oxidative decarboxylation with fragmentation of the intermediate endoperoxide afforded the isochrysohermidin subunit **21** in 89% yield.[14c] In initial efforts, the $^1O_2$ was generated photochemically in the presence of the photosensitizer, Rose Bengal. However, since the photosensitizer was difficult to completely remove from the reaction mixtures, a resin-bound form of Rose Bengal was used, which was found to effect the desired transformation without any decrease in reactivity or product yields.[15]

With the isochrysohermidin subunit **21** in hand, the hybrid conjugates were prepared by coupling with the distamycin substructures incorporating either two or three *N*-methyl pyrrole subunits. Accessing isochrysohermidin–dipyrrole analogue **3** was achieved by treatment of **10** with **21** in the presence of EDCI/DMAP to provide **3** in 31% yield. Similarly, treatment of dipyrrole **13** followed by addition of **21** in the presence of EDCI/DMAP provided **4** in 32% yield after purification. The tripyrrole conjugates **5** and **6** were prepared from the corresponding tripyrroles **12** and **15**, respectively. The tripyrrole conjugate **5** was prepared from coupling **12** and **21** in the presence of EDCI and *i*-Pr_2NEt to provide **5** in 37% yield after purification by column chromatography. Similarly, the tripyrrole conjugate incorporating the *N*,*N*-dimethylpropylamine tail (**6**) was obtained from EDCI and *i*-Pr_2NEt mediated coupling of **15** and **21** to provide **6** in 32% yield (Scheme 4).

DNA Binding Affinity. The DNA binding properties of compounds **3**–**6** were first established by using a fluorescence intercalator displacement (FID) assay titration to establish a binding constant (*K*).[16] This method is based on the loss of fluorescence derived from the titration displacement of ethidium bromide from a DNA of interest. The agents were examined for their ability to bind a hairpin deoxyoligonucleotide containing a central five base pair AT-rich binding site (AAAAA) adjacent to capping GC base pairs relative to distamycin A and results are summarized in Table 1.

TABLE 1. DNA Binding Affinity of 3–6 Compared to Distamycin and Its Derivatives

Agent	Poly-d[A]–poly-d[T] K ($\times 10^6$ M^{-1})	5′-CGAAAAAC A A A 3′-GCTTTTTG A A A K ($\times 10^6$ M^{-1})
distamycin A	15.0	17.0
22	15.9	15.5
23	2.1	6.4
3	--	6.3
4	--	5.9
5	--	8.9
6	--	13.8

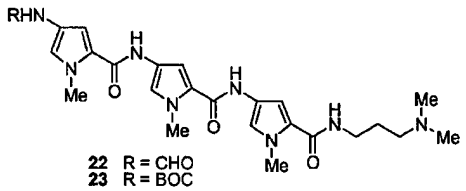


FIGURE 2. Distamycin analogues utilized in the FID assay.

Additionally, the binding affinities of several other distamycin derivatives (**22** and **23**, Figure 2) are included for comparison. Distamycin A binds to poly-d[A]-poly-d[T] and the 5′-AAAAA-3′ hairpin deoxyoligonucleotide with essentially the same affinity. Moreover, replacement of the amidine side chain on distamycin with *N*,*N*-dimethylpropylamine (**22**) simplifies the synthesis and does not adversely affect binding affinity. By contrast, substitution at the *N*-terminus of distamycin analogues has more of an impact on DNA binding affinity. Replacement of the *N*-formyl group with a sterically bulky BOC group with **23** lowers the binding affinity and suggests that large substituents at the *N*-terminus are not as well accommodated in the minor groove. Dipyrrole hybrids **3** and **4** show a 3-fold decrease in binding affinity for the hairpin deoxyoligonucleotide. These derivatives possess one less pyrrole subunit than distamycin and are therefore expected to be less effective noncovalent DNA binding agents. Interestingly, both **3** and **4** exhibited an affinity greater than expected and there is essentially no difference in binding affinity between **3** and **4** potentially representative of a DNA alkylation event. Tripyrrole derivative **5** lacks the C-terminal basic side chain and has half the binding affinity of the natural product. However, by incorporating three *N*-methylpyrrole subunits as well as the basic side chain into the hybrid **6**, it is possible to obtain a binding affinity close to that of distamycin and its closest analogue **22**. Disappointingly, **3**–**6** exhibited no time-dependent increase in binding affinity indicative of a slow, reversible covalent attachment to DNA. Thus, although the surprisingly effective

(15) Bernasconi, C.; Cottier, L.; Descotes, G.; Nigay, H.; Pardon, J. C.; Wisniewski, A. *Bull. Soc. Chim. Fr.* **1984**, *7*, II-323.

(16) (a) Boger, D. L.; Fink, B. E.; Brunette, S. R.; Tse, W. C.; Hedrick, M. P. *J. Am. Chem. Soc.* **2001**, *123*, 5878. (b) Boger, D. L.; Tse, W. C. *Bioorg. Med. Chem.* **2001**, *9*, 2511.

behavior of **5** might suggest a covalent attachment to DNA, the behavior of **6** relative to **22** along with the lack of time-dependent binding affinity (data not shown) suggests it is not observed. Although it is possible that the covalent attachment is rapidly reversible, the intrinsic stability of the carbinolamide of isochrysohermidin ($t_{1/2}$ ca. 24−48 h, DMSO) suggests that is also unlikely. Thus, although we do not yet have a good explanation for the surprising behavior of **3**−**5**, we are confident that it is not derived from a stable, slowly reversible covalent attachment to DNA.[17]

Experimental Section

Dimethyl 3-Methoxy-1-[(3-methoxycarbonyl)propyl]-1*H*-pyrrole-2,5-dicarboxylate (19). Methyl 4-bromobutyrate (309 μL, 2.44 mmol) was added to a solution of **18**[14] (281 mg, 1.32 mmol) and K_2CO_3 (455 mg, 3.29 mmol) in anhydrous DMF (20 mL). The reaction mixture was warmed at 80 °C and stirred under N_2. After 4 h, the reaction mixture was cooled to 25 °C, poured into H_2O (100 mL), and extracted with CH_2Cl_2 (4 × 100 mL). The combined organic layers were dried (Na_2SO_4), filtered, and concentrated under vacuum. Chromatography (SiO_2, 50% EtOAc−hexanes) afforded **19** (399 mg, 96%) as a white solid: mp 74−75 °C; ^{1}H NMR ($CDCl_3$, 400 MHz) δ 6.52 (1H, s), 4.81 (2H, t, J = 7.3 Hz), 3.85 (3H, s), 3.82 (6H, s), 3.64 (3H, s), 2.31 (2H, t, J = 7.6 Hz), 2.04 (2H, quint, J = 7.9 Hz); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 173.2, 161.1, 160.6, 152.5, 124.1, 112.8, 101.0, 57.9, 51.7, 51.5, 51.4, 45.4, 31.0, 26.7; MALDI-HRFTMS *m/z* 336.1056 (M + Na^+, $C_{14}H_{19}NO_7$ requires 336.1054).

1-(4-Butyric acid)-4-methoxy-5-methoxycarbonyl-1*H*-pyrrole-2-carboxylic Acid (20). $LiOH·H_2O$ (117 mg, 2.80 mmol) was added to a stirred solution of **19** (399 mg, 1.27 mmol) in a 2:1:1 solution of THF:MeOH:H_2O (8 mL). After 20 h, the mixture was partitioned between Et_2O and H_2O. The aqueous layer was acidified with the addition of 5% aqueous HCl (pH 3.0) and extracted with EtOAc (4 × 20 mL). The combined EtOAc layers were dried (Na_2SO_4), filtered, and concentrated under vacuum to afford diacid **20** (358 mg, 99%) as a white solid: mp 193−195 °C; ^{1}H NMR ($CDCl_3$, 400 MHz) δ 6.61 (1H, s), 4.79 (2H, t, J = 7.3 Hz), 3.81 (3H, s), 3.79 (3H, s), 2.25 (2H, t, J = 7.4 Hz), 2.04 (2H, quint, J = 7.9 Hz); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 176.7, 163.1, 162.8, 154.2, 126.3, 113.8, 102.5, 58.3, 51.6, 46.3, 31.9, 28.1; MALDI-HRFTMS *m/z* 308.0744 (M + Na^+, $C_{12}H_{15}NO_7$ requires 308.0741).

Isochrysohermidin Subunit 21. A 3:1 solution of CH_3CN−H_2O (40 mL) was added to a quartz flask charged with **20** (23 mg, 0.08 mmol) and Rose Bengal resin (7.0 mg, 0.0006 mmol).[18] The solution was irradiated under a Hanovia high-pressure mercury lamp (450 W) through a uranium yellow glass filter (transmits <330 nm) with a steady stream of O_2 bubbled through the solution. After 3 h, a small amount of activated charcoal was added to remove any solubilized Rose Bengal and the solution filtered through Celite and rinsed with MeOH. The solvent was concentrated under vacuum to afford **21** (21.8 mg, 86% yield) as a transparent glass: ^{1}H NMR ($CDCl_3$, 400 MHz) δ 5.14 (1H, s), 4.62 (1H, s), 3.79 (3H, s), 3.77 (3H, s), 3.68 (1H, dt, J = 14.4, 6.6 Hz), 3.08 (1H, dt, J = 14.5, 6.6 Hz), 2.32 (2H, t, J = 7.2 Hz), 1.79 (2H, quint, J = 7.3 Hz); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 172.8, 170.9, 170.8, 167.3, 94.8, 63.6, 59.0, 53.4, 40.2, 25.2, 23.6; MALDI-HRFTMS *m/z* 296.0728 (M + Na^+, $C_{11}H_{15}NO_7$ requires 296.0741).

Compound 3. 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDCI) (33 mg, 0.17 mmol) and DMAP (11 mg, 0.08 mmol) were added to a mixture of **21** (12 mg, 0.04 mmol) and **10** (29 mg, 0.08 mmol) in anhydrous DMF (0.5 mL) and the reaction mixture was stirred at 0 °C under N_2. After 16 h, the reaction mixture was diluted with 1:1 *i*-PrOH−$CHCl_3$ (10 mL) and washed with 10% aqueous $NaHCO_3$ (2 × 20 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated under vacuum. Chromatography (SiO_2, 15:1 $CHCl_3$–MeOH) afforded **3** (3.3 mg, 31%) as an off-white syrup: ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.84 (1H, br s), 7.73 (1H, s), 7.41 (1H, s), 7.15 (1H, s), 6.77 (1H, s), 6.67 (1H s), 5.06 (1H, s), 3.90 (3H, s), 3.89 (3H, s), 3.85 (3H, s), 3.82 (3H, s), 3.81 (3H, s), 3.57 (1H), 2.36 (2H, t, J = 6.4 Hz); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 175.6, 172.6, 165.0, 163.2, 161.5, 124.6, 123.9, 123.4, 122.6, 121.0, 110.5, 106.1, 94.2, 59.8, 54.1, 51.6, 39.4, 37.1, 37.0, 34.7, 26.5; MALDI-HRFTMS *m/z* 554.1862 (M + Na^+, $C_{24}H_{29}N_5O_9Na$ requires 554.1857).

Compound 4. EDCI (29 mg, 0.151 mmol) was added to a mixture of **21** (10 mg, 0.038 mmol), **13** (29 mg, 0.076 mmol), and DMAP (18 mg, 0.151 mmol) in anhydrous DMF (0.3 mL) and the reaction mixture was stirred at 25 °C under N_2. After 24 h, the reaction mixture was diluted with 1:1 *i*-PrOH−$CHCl_3$ (10 mL) and washed with H_2O (2 × 10 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated under vacuum. Chromatography (RPC_{18}-PTLC, 4:1 MeOH−50 mM HCO_2NH_4 buffer) afforded **4** (7 mg, 32%) as a clear syrup: ^{1}H NMR (2:1 CD_3OD−CH_2Cl_2, 500 MHz) δ 8.44 (1H, s), 7.15 (1H, d, J = 1.8 Hz), 7.11 (1H, d, J = 1.8 Hz), 6.87 (1H, d, J = 1.8 Hz), 6.83 (1H, d, J = 1.8 Hz), 5.18 (1H, s), 3.88 (3H, s), 3.87 (3H, s), 3.85 (3H, s), 3.78 (3H, s), 3.40 (4H, m), 3.20 (3H, m), 2.88, (6H, s), 2.32 (2H, t, J = 7.7 Hz), 1.97 (3H, m); ^{13}C NMR (2:1 CD_3OD−CH_2Cl_2, 125 MHz) δ 177.0, 176.9, 172.5, 159.6, 126.9, 123.3, 120.8, 120.5, 106.5, 105.9, 99.8, 94.0, 59.7, 56.6, 55.5, 43.7, 39.3, 36.6, 34.5, 30.7, 27.4, 27.3, 26.6, 26.3; MALDI-HRFTMS *m/z* 602.2937 (M + H^+, $C_{28}H_{39}N_7O_8$ requires 602.2933).

Compound 5. EDCI (35 mg, 0.18 mmol) was added to a mixture of **21** (12 mg, 0.04 mmol), **12** (39 mg, 0.09 mmol), and *i*-Pr_2NEt (16 μL, 0.09 mmol) in anhydrous DMF (0.5 mL). The reaction mixture was stirred under N_2 at 0 °C for 3 h and allowed to warm to 25 °C. After 18 h, the reaction mixture was diluted with 1:1 *i*-PrOH−$CHCl_3$ (10 mL) and washed with H_2O (2 × 10 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated under vacuum. Chromatography (SiO_2, 12:1 $CHCl_3$−MeOH) afforded **5** (11 mg, 37%) as an off-white syrup: ^{1}H NMR (1:1 CD_3OD−CD_2Cl_2, 500 MHz) δ 7.68 (1H, s), 7.34 (1H, d, J = 2.2 Hz), 7.18 (1H, d, J = 2.2 Hz), 7.11 (1H, d, J = 1.8 Hz), 6.91 (1H, d, J = 2.2 Hz), 6.87 (1H, d, J = 2.2 Hz), 6.80 (1H, d, J = 1.8 Hz), 5.12 (1H, s), 3.89 (3H, s), 3.88 (3H, s), 3.87 (3H, s), 3.84 (3H, s), 3.78 (3H,s), 3.77 (3H, s), 3.45 (1H, dt, J = 14.5, 7.0 Hz), 3.16 (1H, dt, J = 14.5, 7.0 Hz), 2.32 (2H, t, J = 7.5 Hz), 1.86 (2H, m); ^{13}C NMR (1:1, CD_3OD−CD_2Cl_2, 125 MHz) δ 174.7, 173.3, 172.0, 169.1, 162.7, 160.8, 124.2, 124.1, 123.2, 122.8, 122.6, 122.0, 120.5, 120.3, 120.1, 109.8, 105.8, 105.3, 93.7, 88.6, 78.8, 59.4, 51.4, 38.9, 36.9, 36.7, 34.1, 30.3, 26.0; MALDI-HRFTMS *m/z* 653.2459 (M^+, $C_{30}H_{35}N_7O_{10}$ requires 653.2440).

Compound 6. EDCI (31 mg, 0.180 mmol) was added to a mixture of **21** (11 mg, 0.04 mmol), **15** (39 mg, 0.08 mmol), and *i*-Pr_2NEt (14 μL, 0.090 mmol) in anhydrous DMF (0.5 mL). The reaction mixture was stirred under N_2 at 0 °C for 3 h and allowed to warm to 25 °C. After 24 h, the reaction mixture was diluted with 1:1 *i*-PrOH−$CHCl_3$ (10 mL) and washed with 10% aqueous $NaHCO_3$ (2 × 10 mL). The organic phase was dried (Na_2SO_4), filtered, and concentrated under vacuum. Chromatography (RPC_{18}-PTLC, 6:1 MeOH−50 mM HCO_2NH_4 buffer) afforded **4** (7 mg, 32%) as a clear syrup: ^{1}H NMR (2:1 CD_3OD−CD_2Cl_2, 500 MHz) δ 7.17 (1H, br s), 7.09 (1H, s), 6.92 (1H, s), 6.86 (1H, s), 6.91 (1H, br s), 6.82 (1H, s), 6.79 (1H, s), 5.09 (1H, s), 3.90 (3H, s), 3.88 (12H, br s), 3.39 (4H, br t, J = 6.3 Hz), 3.13 (2H, m), 3.05 (2H, br t, J = 7.7 Hz), 2.80 (6H, br s), 2.31 (2H, m), 1.96 (2H, m), 1.83 (2H, m), 0.877 (2H, br t, J = 6.6 Hz); ^{13}C NMR (1:1, CD_3OD−CD_2Cl_2, 125 MHz) δ 173.0, 170.3, 169.4, 168.1, 167.0, 164.9, 161.4, 124.6, 123.8, 123.4,

(17) Cytotoxic activity: L1210 IC_{50} = 67 (**1**), 119 (**3**), 165 (**4**), >200 (**5**), and >200 μM (**6**).

(18) Concentration of resin-bound Rose Bengal was determined according to ref 15 (0.09 mol of Rose Bengal/g of support).

123.3, 120.9, 120.7, 120.5, 106.5, 106.4, 105.9, 56.7, 47.4, 47.3, 43.8, 41.1, 37.1, 37.0, 36.7, 34.4, 30.7, 27.4, 27.3, 26.6, 25.3, 23.7, 23.1; MALDI-HRFTMS *m/z* 723.3337 ($M + H^+$, $C_{34}H_{45}N_9O_9$ requires 723.3340).

Determination of DNA Binding Constants. A 3-mL quartz cuvette was loaded with Tris buffer (0.1 M Tris, 0.1 M NaCl, pH 8) and ethidium bromide (0.44×10^{-5} M final concentration). The fluorescence was measured (excitation 545 nm, emission 595 nm, EtBr) and normalized to 0% relative fluorescence. The 5′-AAAAA-3′ hairpin deoxyoligonucleotide was added (1.5 μM, 12 μM in base pair final concentration), and the fluorescence measured again and normalized to 100% relative fluorescence. A solution of the agent (3 μL, 0.1 mM in DMSO) was added, and the fluorescence measured following 5 min of incubation at 23 °C. Subsequent addition of 3-μL aliquots of the agent was continued until the system reached saturation and the fluorescence remained constant with successive compound additions.

Scatchard Analysis of the Titration Curve. The ΔF was plotted versus molar equivalents of agent and the ΔF_{sat} was determined mathematically by solving the simultaneous equations representing the pre- and postsaturation regions of the titration curve. Utilizing eqs 1–3, a Scatchard plot was generated where ΔF/[free agent] was plotted vs ΔF. The slope of the region immediately preceding complete saturation of the system provided $-K$.[16]

$$\left(\frac{\Delta F_x}{\Delta F_{sat}}\right)\frac{1}{X} = \text{fraction of DNA} - \text{agent complex} \quad (1)$$

$$\left[1 - \left(\frac{\Delta F_x}{\Delta F_{sat}}\right)\frac{1}{X}\right] = \text{fraction of free agent} \quad (2)$$

$$[\text{DNA}]_T\left[X - \frac{\Delta F_x}{\Delta F_{sat}}\right] = [\text{free agent}] \quad (3)$$

Acknowledgment. We gratefully acknowledge the financial support of the National Institutes of Health (CA 41986, CA 42056) and the Department of the Army (DAMD17–02–1–0135).

Supporting Information Available: ^{1}H NMR of all new compounds **3**–**6** and **19**–**21**. This material is available free of charge via the Internet at http://pubs.acs.org.

JO034326C

A Fluorescent Intercalator Displacement Assay for Establishing DNA Binding Selectivity and Affinity

WINSTON C. TSE AND DALE L. BOGER*
Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037

Received May 13, 2003

ABSTRACT
A summary of the qualitative and quantitative elements of a fluorescent intercalator displacement (FID) assay useful for establishing the DNA binding selectivity, affinity, stoichiometry, and binding site size and distinguishing modes of DNA binding is provided.

Introduction

The regulation of gene expression is based on the sequence-selective recognition of nucleic acids by repressor, activator, and enhancer proteins. A full understanding of the proteins involved, the delineation of the sequences to which they bind, and the discovery of the genes that they regulate holds significant promise in therapeutic medicine.[1-6] Thus, extensive efforts continue to be directed at understanding the transcriptional process and are being increasingly directed at the discovery of small molecules that selectively bind DNA and activate (block a repressor) or inhibit (block an activator) gene expression.[7-11]

Of the techniques commonly used to establish the DNA binding properties of small molecules and proteins, most are technically challenging, require the knowledge of specialized biochemical procedures, and are time and labor intensive. The most widely used methods are footprinting[12] and affinity cleavage.[13] Because of the power of the technique, a number of such methods have been introduced including DNase I,[14] exonuclease III,[15] MPE–Fe(II),[16] 1,10-phenanthroline–Cu(I),[17] and EDTA–Fe(II)[18] footprinting. Complementary approaches for disrupting binding (interference footprinting)[19] by specific base or phosphate modifications have also been introduced. Less general techniques that capitalize on a compound's intrinsic DNA cleavage (e.g., bleomycin, endiynes), alkylation/thermal cleavage[20] (e.g., CC-1065, duocarmycins), alkylation (inhibition of in vitro transcription), or cross-linking[21,22] (e.g., mitomycin) properties have been applied to selected classes of molecules. Recently, techniques for expanding the sequence coverage of footprinting[23] and the use of electrophoretic mobility shift assays (EMSA, gel retardation assay)[24] and DNase I footprinting[25,26] for the iterative deconvolution of mixture libraries have been disclosed, expanding their applications. Inherent in these methods is the characterization of the highest affinity sites within a size-limited segment of DNA. Similarly, the DNA binding properties of proteins[27] are typically assessed by selection screening,[28-30] footprinting,[14] or EMSA.[31] The former provides exhaustive sequence coverage for deducing the preferred site(s), but it selects only the highest affinity sites and does not provide quantitative binding information. Footprinting and EMSA[31] have been used to define, or at least refine, a protein's binding selectivity, but their most frequent uses have been to provide qualitative distinctions and quantitative comparisons among candidate binding sites or those constructed to assess single base pair (bp) substitutions.

Herein, we review a complementary technique, a fluorescent intercalator displacement (FID) assay,[32,33] for establishing DNA binding affinity, sequence selectivity, and binding stoichiometry. The assay is nondestructive, technically nondemanding, and amenable to high-throughput screening. The former feature would permit DNA immobilization onto reusable supports for repetitive use and expansion of the sequence space beyond that presently exemplified (all 5-bp sites). For a single compound, the technique permits establishment of a rank order binding profile for all possible 5-bp sites, comprehensively defining the sequence selectivity in a single experiment. For a defined sequence, it permits the high-throughput identification of binding agents from a library of compounds or quantitative titrations for establishment of binding constants. The assay is not limited to small molecule assessments and has been used with a variety of ligands, including proteins and triplex-forming oligonucleotides.

The FID Assay

The assay utilizes the displacement of ethidium bromide (or thiazole orange)[34,35] from hairpin deoxyoligonucleotides (Figure 1). Hairpin DNAs are treated with the intercalator, yielding a fluorescence increase upon binding. Addition of a DNA binding compound results in a decrease in fluorescence due to displacement of the bound intercalator, where the percent fluorescence decrease is directly related to the extent of binding. For a panel of DNA sequences displayed in a 96-well format, the resulting profile of percent fluorescence decrease provides relative binding affinities and comprehensively defines a compound's sequence selectivity. With individual sequences, quantitative titration of a compound against a hairpin prebound with ethidium provides reliable binding constants and the stoichiometry of binding. These two systems may be used in a complementary fashion, one intended for a high-throughput screen

Dale L. Boger is the Richard and Alice Cramer Professor of Chemistry at The Scripps Research Institute and a member of the Skaggs Institute for Chemical Biology.

Winston C. Tse received a B.A. from Northwestern University, where he worked in the laboratory of Professor Joseph B. Lambert, and his Ph.D. from The Scripps Research Institute with Professor Boger. He joined Gilead Sciences in 2003.

* Corresponding author. Tel.: (858) 784-7522. Fax: (858) 784-7550. E-mail: boger@scripps.edu.

10.1021/ar030113y CCC: $25.00
Published on Web 00/00/0000

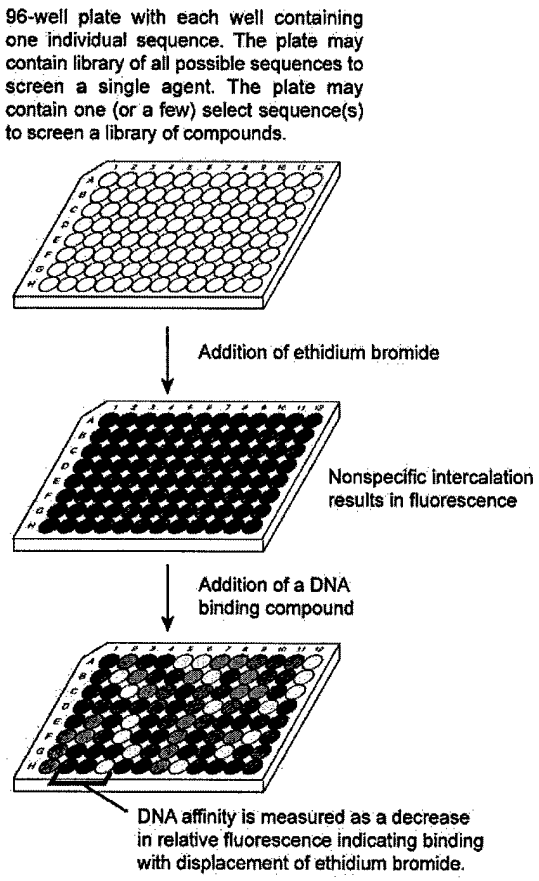


FIGURE 1. The FID assay.

capable of handling large compound libraries or providing comprehensive sequence preference data, and one intended for quantitative elucidation of DNA–ligand interactions and establishing binding constants.

Following the characterization of the intrinsic fluorescent increase that accompanies DNA intercalation,[36] the displacement of DNA-bound ethidium bromide has been widely used to establish DNA binding.[37] Ethidium's nonselective,[38] rapidly equilibrating, and low binding affinity ($\sim 10^5$ M^{-1})[36] allows assessment of compounds with low binding affinity or assessment of low affinity sites, aspects not easily addressed by other techniques. For tight binding sequences, the ethidium displacement proceeds in a virtually noncompetitive manner, permitting quantitative binding assessments.

Thiazole orange[34,35] is an effective alternative intercalator that addresses three issues: (1) its excitation and emission maxima are distinct from those of ethidium, (2) its fluorescent enhancement upon intercalation exceeds that of ethidium (ca. 3000-fold vs 20-fold), and (3) it displays less sequence-dependent DNA binding, albeit with a higher affinity.[39] This permits the examination of systems where there would be ethidium fluorescence interference and screening at lower concentrations or with greater sensitivities than achievable with ethidium, and

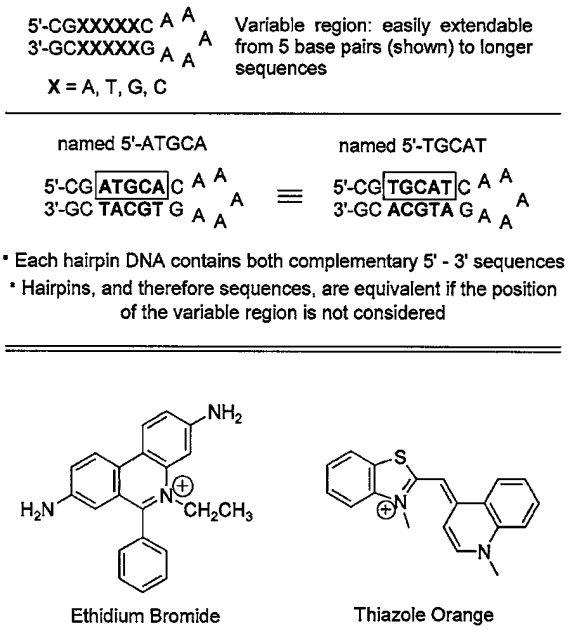


FIGURE 2. Structures of hairpin deoxyoligonucleotides, ethidium bromide, and thiazole orange.

with less sequence variation and greater consistency in the absolute fluorescence readings. However, the binding constants established with thiazole orange are slightly lower than those established with ethidium.[32,39] This reflects the higher association constant and greater competitive binding of thiazole orange, reducing the measured apparent binding constant.

Hairpin deoxyoligonucleotides[40] (Figure 2) proved especially useful in the FID assay. Embedded in the hairpin are two complementary 5′-to-3′ sequences, connected by a loop, avoiding the requirement for two separate strands and the associated additional quantitation and mixing. The number of hairpins required to create a library of sequences is half the number of sequences. For example, 512 hairpins are required for a library of all 1024 possible 5-bp sequences (Figure 2). Moreover, the hairpins were established to provide stable duplexes at working temperatures (25 °C), independent of the sequence.[32,33] The variable most critical to the success of the assay, and most likely to be responsible for avoidable errors, is the quality of the hairpins. In addition to the concern about their constitution and purity, their concentration is critical and may be determined by measuring the UV absorption (260 nm) of the denatured, single-stranded DNA at 80–95 °C. Since the hairpins exist in a construct representing a combination of double- (stem) and single-strand (loop) DNA at 25 °C, calculations based on the UV absorption at 25 °C, utilizing the standard coefficients useful for single-stranded deoxyoligonucleotides, underestimate the concentration by as much as 25%.[41]

Binding Constants and Stoichiometry

Quantitative displacement of ethidium from a hairpin deoxyoligonucleotide provides a well-defined titration curve that is useful for establishing binding constants and

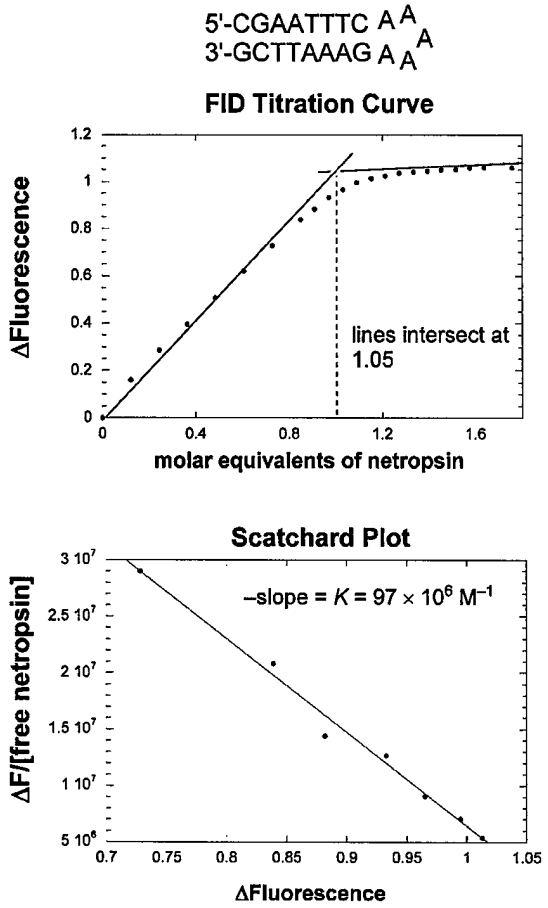


FIGURE 3. FID titration curve of netropsin against a hairpin deoxyoligonucleotide containing the sequence 5'-AATTT, and Scatchard plot for K_a (–slope) determination.

the stoichiometry of binding. A plot of the change in fluorescence versus equivalents of compound provides a titration curve from which the stoichiometry of binding may be derived as the intersection of the pre- and postsaturation portions of the curve (Figure 3). This method, analogous to that introduced by Bruice with Hoechst 33258,[42,43] is easily extended to analyzing higher order 2:1 and 3:1 complexes. For 1:1 binding, binding constants are established by Scatchard analysis[44] of the equilibrium portion of the titration curve, generating a plot of ΔF/[free agent] versus ΔF, yielding a linear section where the slope provides K_a. Binding constants produced by this indirect technique involving the displacement of ethidium yielded results comparable to those obtained directly by monitoring the fluorescent increase of selected fluorescent DNA binding compounds (e.g., DAPI, Table 1).

DNA Binding Sequence Selectivity

Minor Groove Binding Compounds. In developmental stages of the assay, distamycin, netropsin, DAPI, Hoechst 33258, and berenil were examined (Figure 4).[32] DAPI, Hoechst 33258, and berenil are fluorescent dyes, enabling a direct assessment of binding. Notably, the fluorescence enhancement characteristic of their binding did not interfere with the measurement of the fluorescence decrease derived from ethidium displacement (e.g., for EB, ex. 545 nm, em. 595 nm; for DAPI, ex. 372 nm, em. 454 nm).

Table 1. DAPI Binding Constants

DNA sequence	K ($\times 10^6$ M^{-1})[a]	K ($\times 10^6$ M^{-1})[b]
5'-AATTT	110	120
5'-AATAA	59	87
5'-ATTAA	52	77
5'-AAAAA	50	65

[a] Scatchard analysis of ethidium displacement titration. [b] Direct titration using fluorescence enhancement of DAPI.

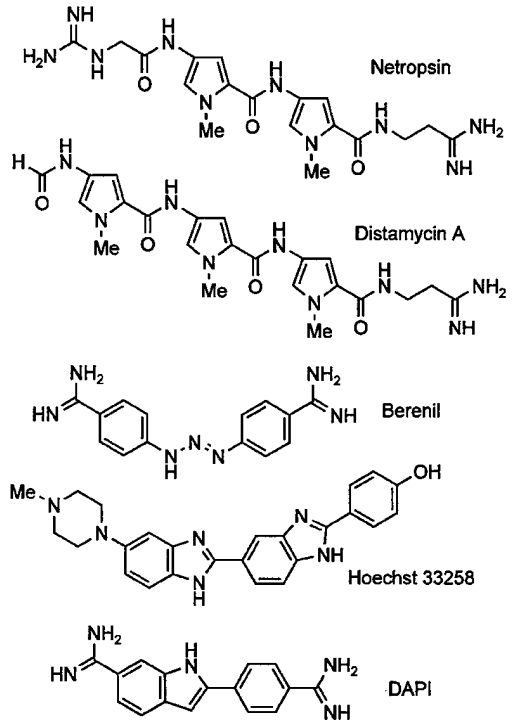


FIGURE 4.

Each compound was screened against a library of 512 DNA hairpins (1.5 μM) containing all possible 5-bp sites in a 96-well format, enlisting three compound concentrations (1.0, 1.5, and 2.0 μM) in duplicate and a Gemini SpectraMax plate reader.[32] On this scale and at these concentrations, the cost of the purchased hairpins is approximately $100/assay. This provided a rank order binding profile for all possible 5-bp sites that is represented as a merged bar graph in Figure 5 for distamycin. In addition to rapidly providing a comprehensive definition of each compound's sequence selectivity, the comparisons provided insights not previously easily recognized. While all compounds displayed the expected A/T binding selectivity, netropsin was the most A/T selective and, by some accounts, distamycin was the least. All exhibit tight binding to 5-bp > 4-bp > 3-bp A/T sites (Figure 6). This preference is related in part to the conformational characteristics of DNA, where longer A/T sites possess a narrower, deeper minor groove known to contribute to the selective binding. Interestingly, distamycin displayed an affinity for GC bp interrupted 5-bp A/T sites (2 × 2-bp) that exceeded even that of 3-bp A/T

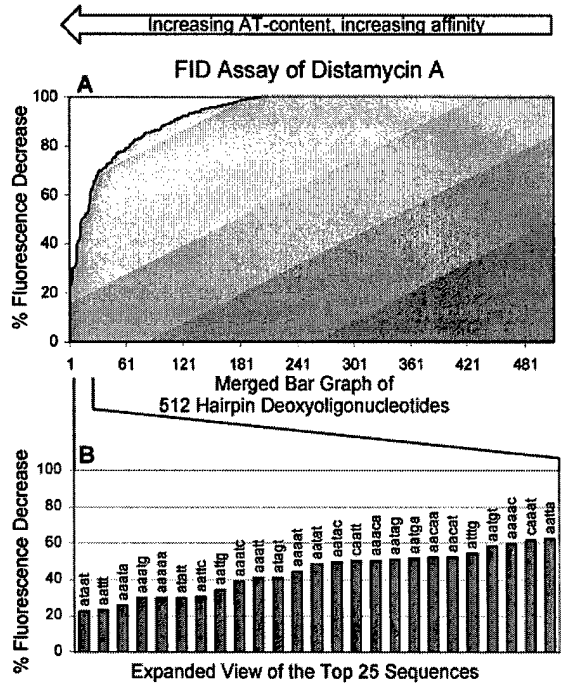


FIGURE 5. FID assay of distamycin A against a library of 512 hairpin deoxyoligonucleotides: (A) merged bar graph; (B) the top (highest affinity) 25 sequences.

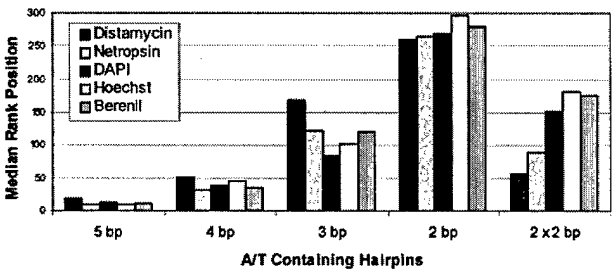


FIGURE 6. Median rank positions of hairpin DNAs containing five, four, three, two, and two × two contiguous A/T bp for the five compounds surveyed.

sites and was only slightly weaker that that of 4-bp A/T sites. Netropsin exhibited a weaker preference for such 2 × 2-bp A/T sites, reduced from its affinity to 4-bp A/T sites. In constrast, DAPI, Hoechst 33258, and berenil exhibited a clear preference for a 3-bp A/T site over a 2 × 2-bp A/T site. Combined, this was suggested to reflect the larger 5-bp binding site requirement for distamycin and its unique compensating ability to bind selected GC bp interrupted 5-bp A/T sites. For each compound, binding constants were determined by ethidium displacement titrations with selected sequences. These were found to be comparable to those established by footprinting or calorimetry, and to be nearly identical (within 2-fold) to those obtained by directly monitoring the fluorescence increase of DAPI and Hoechst 33258 upon DNA binding (see Table 1).[32]

A similar screen of netropsin against a library of all possible 4-bp sites was disclosed, and binding constants for all 10 possible 4-bp A/T sites were determined by titrations (Table 2).[39] With binding constants for all 10

Table 2. Comparative Titration Binding Constants

titrant	DNA sequence	K ($\times 10^6$ M^{-1})	binding stoichiometry
EB	5′-AATT	2.7	2.92
TO	5′-AATT	15	3.05

		K ($\times 10^6$ M^{-1})	
titrant	DNA sequence	ethidium bromide	thiazole orange
netropsin	5′-AAAT	127	113
	5′-AAAA	92	71
	5′-AATT	65	54
	5′-AATA	64	44
	5′-ATTA	45	35
	5′-ATAT	41	33
	5′-ATAA	34	18
	5′-TAAA	26	15
	5′-TTAA	11	8
	5′-TATA	11	8

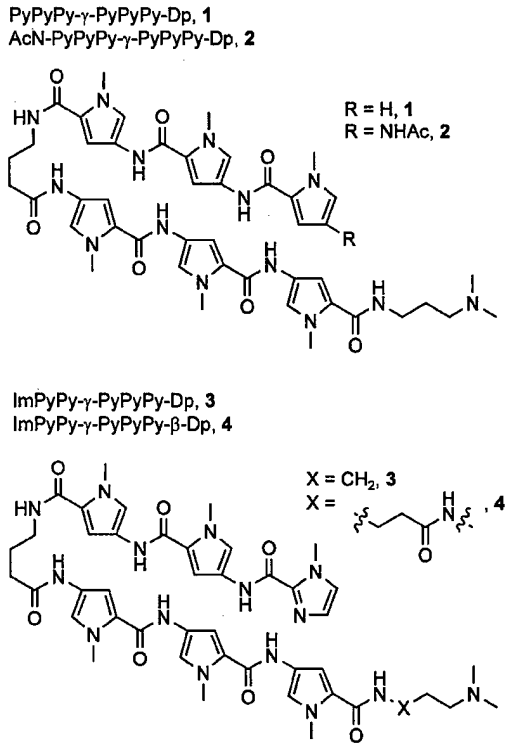


FIGURE 7.

possible A/T sites, an explicit rank order binding was established that was more comprehensive than that provided in all prior combined footprinting studies. Highlighted in these studies was the fact that the assay allows the rapid identification of the preferred ensemble of sequences for a compound, but that substantive conclusions drawn about adjacent sequences in the 96-well screen should be reserved for quantitative titrations.

Hairpin Polyamides. An analysis of four pyrrole (Py)/imidazole (Im) hairpin polyamides (Figure 7) was conducted that not only served to establish the utility of the assay for examining minor groove binding polyamides, but also underscored its comprehensive capabilities. The polyamides were chosen to examine the subtle effects of N- and C-terminal functionalization on the DNA binding selectivity, representing a challenging test of the assay

Table 3. Sequence Selectivity of ImPyPy-γ-PyPyPy-β-Dp

A. FID Assay Analysis

Sequence[a]	1.5 μM agent		2.0 μM agent	
	Avg Rank	Avg Score	Avg Rank	Avg Score
5′-WGWWW	55	0.47	54	0.47
5′-WWGWW	82	0.44	36	0.55
5′-WCWWW	97	0.32	105	0.28
5′-WWWWW	131	0.27	111	0.27

B. Binding Constants by FID Titration

Binding Mode[b]	Sequence	K_a (M^{-1})
	1. 5′-TAACAAT[c]	2.39×10^8
	2. 5′-TAACAAC	1.73×10^8
	3. 5′-TAACACC	1.33×10^7
	4. 5′-GAACAAT	3.31×10^7
	5. 5′-GAACAAC	2.84×10^7
	6. 5′-ATTGTTA	9.87×10^7
	7. 5′-GTTGTTA	9.89×10^7
	8. 5′-GGTGTTA	2.57×10^7
	9. 5′-GTTGTTC	3.30×10^7

[a] W = A or T. [b] ○, Py; ●, Im. [c] Lit.[48] $K = 2.9 \times 10^8$ M^{-1}, footprinting.

capabilities. The complete binding profiles of the four molecules to all possible 5-bp sites were established using the FID assay.[45]

The comparison of **1** and **2** revealed the detrimental effect of the *N*-acetyl substitution with **2**,[46,47] and the hairpin polyamide ImPyPy-γ-PyPyPy-Dp (**3**), like **1**, displayed a straightforward correlation with its expected selectivity for 5′-WGWWW. For ImPyPy-γ-PyPyPy-β-Dp (**4**), the ability to screen complete sequence space resulted in the discovery of an unexpected 5′-WWCWW sequence specificity (Table 3A). Although detailed studies have probed the ligand−DNA interactions of this molecule and its sequence selectivity was established (5′-WGWWW),[48] its explicit binding to 5′-WWCWW had not been described. Use of the FID assay to comprehensively screen DNA provided a complete picture of its selectivity. Strikingly, the quantity and quality of the data generated by the assay also provided the basis for a binding model. This model established that **4** has an optimal binding site of 6 bp (not the expected 5 bp) of the form 5′-WWGWWW, and that the β/Dp tail combination requires two (not one) degenerate A/T base pairs. FID titrations confirmed the model by providing quantitative binding constants (Table 3B).

Also disclosed in this study was the analysis of sequence sets and two scoring methods for their comparison (Table 3A). The first simply calculates the average rank order position of the constituent members of a sequence set, whereas the second calculates an average %*F* decrease of the sequence set relative to the highest affinity sequence. This latter, less obvious scoring procedure may better reflect the relative importance of binding to a sequence set.

Selection of Binding Agents for a Defined Sequence(s) from Compound Libraries

Most of the current screening technologies are sufficiently technically demanding that they are most often used to investigate a few individual compounds. A true attribute of the FID assay is its capabilities for screening libraries of compounds (mixtures or individual).[33,49]

The rapid screening of individual sequences was illustrated with two hairpin deoxyoligonucleotides containing two sequences of the androgen response element (ARE), the 14-bp ARE-consensus, and PSA-ARE-3.[33,49] The emergence of hormone-independent, constituently active androgen receptor dimer is responsible for prostrate cancer relapse resistant to chemotherapeutic treatment. At this stage, competitive inhibition of the androgen receptor dimer DNA binding has therapeutic potential and could arise from small molecules with selective affinity toward the ARE-consensus and PSA-ARE-3. Libraries of distamycin-like compounds were prepared using 11 heteroaromatic subunits in addition to the *N*-methylpyrrole native to distamcyin (Figure 8). Libraries were produced such that the first two subunits (B and C) were fixed, producing 132 mixtures of 10 compounds with variations at the last position (A). Screening the library against the two hairpins using the FID assay revealed that the mixture containing the pyrrole subunit at both the second (B) and third (C) positions gave the largest decrease in fluorescence with the PSA-ARE-3 sequence which contains a 5-bp A/T site. The affinity dropped for the ARE-consensus containing a GC bp at the center of the 5-bp A/T site of this sequence. Deconvolution by screening individual compounds of this mixture afforded the distamycin analogue (**5**) as having the highest affinity, followed closely by an analogue containing a thiophene subunit at the position A (Figure 9). Both compounds exhibited diminished affinity for the ARE-consensus sequence resulting from the intervening GC bp.[33,49]

Two other 10-compound mixtures also bound the PSA-ARE-3 sequence effectively. Deconvolution to individual compounds identified **6** and **7** as tight binders comparable to **5**. Notably, **6** showed a loss of affinity for the ARE-consensus analogous to **5**, but **7** retained equal affinity, making it ideal for binding both the PSA-ARE-3 and ARE-consensus sequences (Figure 9). Impressively, **7** exhibited potent (IC_{50} = 8 nM) and selective (ca. 40-fold) inhibition of androgen receptor-mediated gene transcription in a cell-based reporter assay, albeit requiring liposome delivery of the compound for cell penetration and observation of activity.

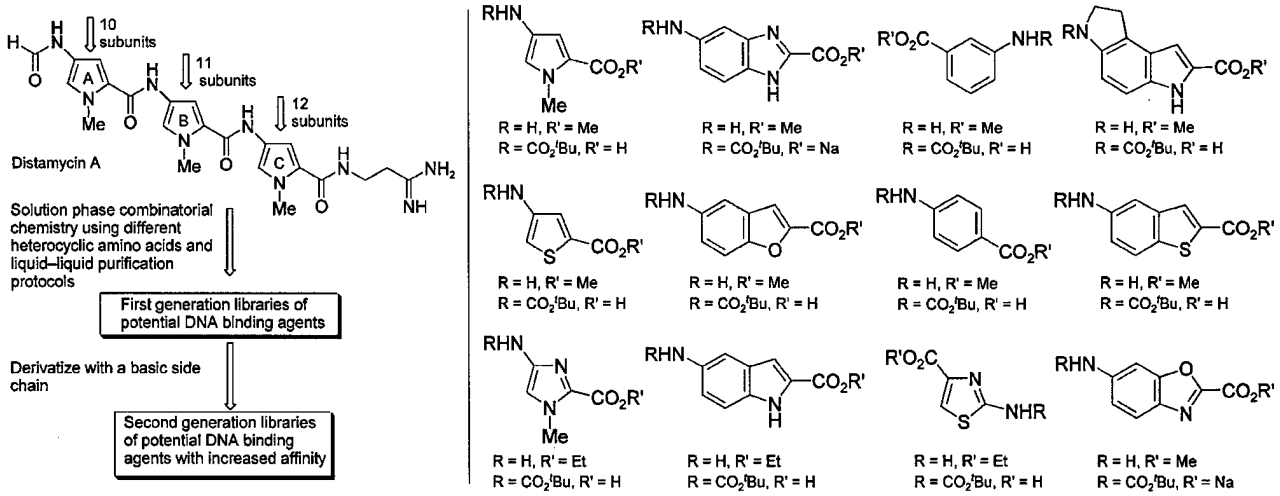


FIGURE 8. Solution-phase strategy for libraries and the subunits used.

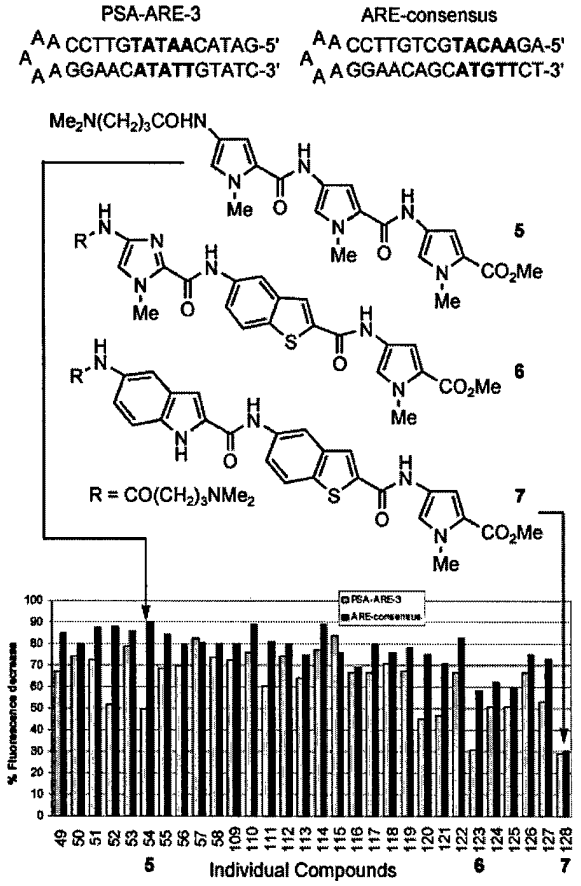


FIGURE 9. FID assay for selecting members of a library with affinity for the ARE-consensus and PSA-ARE-3.

Establishing Subtle Modes of Binding: Hairpin versus Extended Binding of Polyamides

The linkage of polyamides with γ-aminobutyric acid (γ) provides hairpin-bound polyamides that mimic the 2:1 side-by-side antiparallel binding of unlinked polyamides, enhances the binding affinity 10^2-10^4-fold, and improves the binding selectivity. In constrast, polyamides incorporating a one-carbon shorter head-to-tail linker, β-alanine (β), bind in an extended conformation, forming 1:1 or side-by-side antiparallel 2:1 complexes.[50] A variant of the FID titrations was used to study the DNA binding properties of α-substituted β-alanine-linked polyamides and a series of novel iminodiacetic acid (IDA)-linked polyamides (Figure 10).[51,52] In these two series, the bound conformation (hairpin versus extended binding and parallel versus antiparallel binding) could be established by analysis of FID titrations of hairpin deoxyoligonucleotides containing a systematically varied A/T binding site size. Complementary assessments using a combination of footprinting and affinity cleavage techniques[12,14–18] are technically more demanding, require the separate preparation of the linked Fe–EDTA affinity cleavage derivatives, and do not as easily distinguish between such alternative binding modes.

In the first of the two studies, a series of Py polyamides were linked using α-substituted β-alanines to probe their impact on binding.[51] Polyamide binding was assessed by using a series of hairpin deoxyoligonucleotides containing a systematically varied length of the A/T binding site (5–12 bp). The (*R*)-OMe-substituted derivative **9** bound with moderately high affinity, whereas the parent β-alanine **8** bound with intermediate affinity to the shortest hairpin (Figure 10). Very little change in the behaviors of **8** and **9** was seen until the length of the binding site reached 8 bp, where **8** dramatically changed. The binding constant increased by 2 orders of magnitude, and the stoichiometry of binding increased to 2. This corresponds to the expected behavior of extended binding over a 8–9-bp site as an antiparallel 2:1 side-by-side dimer.[53–55] In contrast-,the binding stoichiometry of **9** remained constant at 1:1, displaying a binding constant of $\sim 8 \times 10^7\ M^{-1}$ throughout the range of 5–10 bp until a binding site length of 11 bp was reached, where a second binding event was observed. The behavior seen with the 11- and 12-bp A/T sites represents two sequential binding events of **9** adopting a hairpin conformation, each requiring nonoverlapping 5-bp A/T sites (Figure 10).

The studies illustrated that a generalizable variant of the FID titrations may be utilized to distinguish hairpin versus extended binding, provide information on the

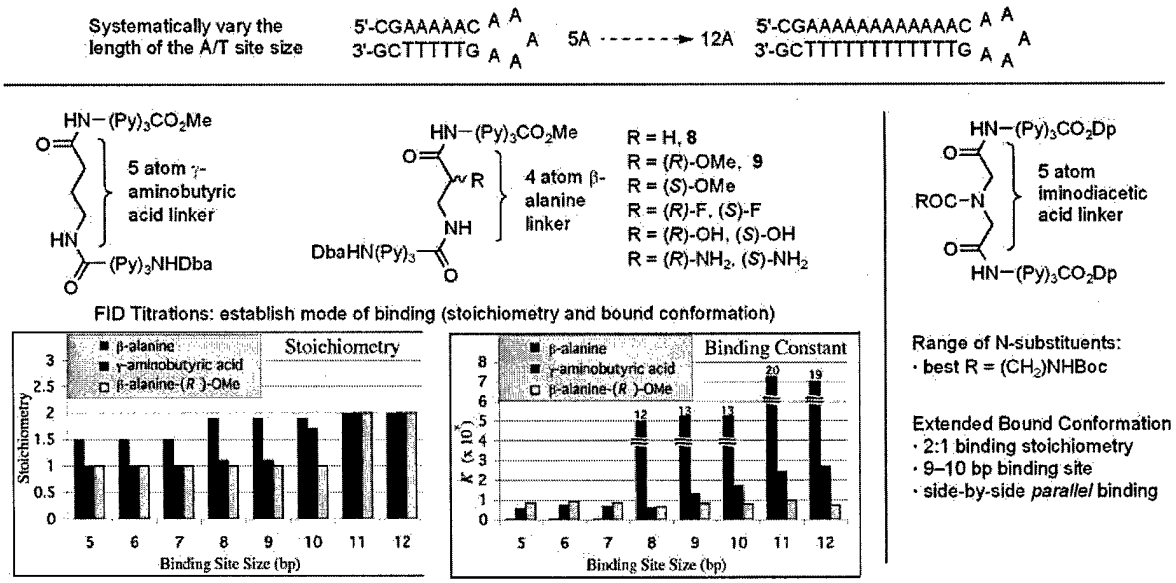


FIGURE 10. Establishment of DNA binding mode for γ-aminobutyric acid-linked, β-alanine-linked, and iminodiacetic acid-linked polyamides. Hairpin deoxyoligonucleotides containing varying A/T site sizes are utilized to determine binding constants and stoichiometry.

binding site size and stoichiometry, and establish absolute affinity. Its use resulted in the discovery that, while most substituents on a β-alanine linker disrupt DNA binding favoring an extended binding mode, (*R*)-α-methoxy-β-alanine maintains strong binding affinity and prefentially adopts a hairpin versus extended binding mode.[51]

In a second disclosure, the DNA binding properties of iminodiacetic acid-linked polyamides were examined (Figure 10). Utilizing the same series of hairpin deoxyoligonucleotides containing a systematically varied A/T binding site size, the binding affinity, stoichiometry, and binding site size were determined. Collectively it was determined that the IDA linker has a unique effect on the ability of the polyamides to bind DNA, providing what appears to be the first well-characterized example of cooperative parallel extended 2:1 side-by-side binding.[52]

DNA Binding Properties of Proteins

The DNA binding affinity and selectivity of proteins[27] are typically assessed by selection screening,[28–30] footprinting and affinity cleavage,[14] or EMSA.[31] Each technique possesses unique strengths, and the FID assay presents a complementary, technically nondemanding method for qualitative or quantitative assessment of DNA bindingwith concurrent establishment of the stoichiometry of binding.

Lymphoid enhancer-binding factor[56] (LEF-1, also referred to as TCF-1α[57,58]) and closely related T-cell factors (TCF-1[59]) are cell type-specific DNA binding proteins that play important regulatory roles.[60,61] The majority of colorectal tumors contain mutations that result in accumulation of β-catenin. β-Catenin binds to and activates transcription factors including LEF-1, which binds the DNA minor groove through a high-mobility-group (HMG) domain, recognizing a consensus sequence 5′-CTTTGWW (W = A or T).[56–63] A recent NMR structure of the complex of the LEF-1 HMG domain and a 15-bp deoxyoligonucleotide[64] revealed that the protein makes extensive and continuous contacts in the DNA minor groove, encompassing the entire region implicated by chemical footprinting and mutagenesis.[56–63]

FID titration of the HMG domain of LEF-1 against a hairpin deoxyoligonucleotide containing 5′-CTTTGAAG provided a well-defined titration curve (Figure 11). Scatchard analysis provided a K_a of 1.5×10^9 M^{-1}, virtually identical to that established by EMSA.[63] Once the viability of monitoring protein−DNA binding with the FID assay was confirmed, several other sequences were rapidly evaluated, including those implicated in initial studies, but not subsequently examined in detail. The third base of the consensus sequence, 5′-CTTTGWW, was permuted to each base variation and revealed that the C, G, and A substitutions bind the LEF-1 HMG domain effectively and almost indistinguishably, exhibiting K_a's less than 2-fold lower than that of the reported consensus sequence (Table 4). As such, the consensus sequence of 5′-CTTTGWW is accurate but could easily by refined to 5′-CTNTGWW (N = G, C, A or T) to reflect this nearly indiscriminant third site.[65]

Although not demonstrated in this work, the FID assay could easily be extended to the examination of a library of proteins or mutants against such hairpins.

Summary and Outlook

The fluorescent intercalator displacement (FID) assay is a rapid, high-resolution, and technically nondemanding technique for establishing DNA binding selectivity and affinity for small molecules, proteins, and oligonucleotides.[66] In a 96-well format, the assay provides for the high-throughput evaluation of a single compound against a library of DNA sequences (establish sequence selectivity) or for the high-throughput selection of high-affinity binders for a defined sequence from a library of compounds.

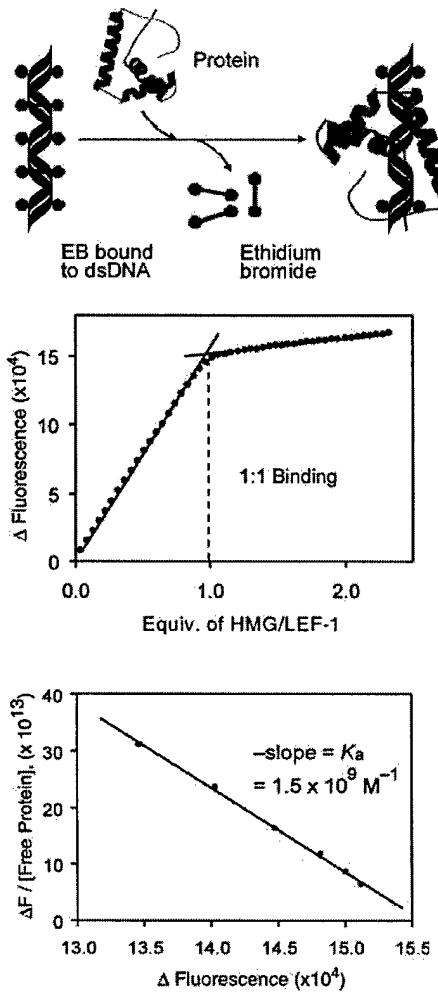


FIGURE 11. Titration of the HMG domain of LEF-1 versus a hairpin DNA containing 5′-CTTTGAA.

Table 4. HMG/LEF-1 Binding Constants

DNA sequence	K_a (M^{-1})
5′-CTTTGAAG[a]	1.5×10^9
5′-CTTTGTTG	1.0×10^9
5′-CCCCGAAG	$\leq 1.0 \times 10^8$
5′-CTCTGAAG	9.4×10^8
5′-CTGTGAAG	8.9×10^8
5′-CTATGAAG	8.3×10^8

[a] Lit.[63] $K_d = 1.0 \times 10^{-9}$ M, EMSA.

Alternatively, FID titrations provide detailed information on single compounds and their binding to individual sequences, including binding constants, stoichiometry of binding, and binding site size. This latter application, as well as the selection screening against a single sequence, is amenable to examination of any sequence length. Its use in screening against a complete library of individual sequences has been exemplified with sequence sizes of ≤5-bp variable regions requiring a library of up to 512 hairpin deoxyoligonucleotides. Longer sequences require increasingly larger numbers of hairpins (e.g., 6 bp =1024, 7 bp = 2048) or a pooling strategy for simple laboratory implementation. Unlike complementary techniques, the FID assay is nondestructive, providing the opportunity for hairpin immobilization onto reusable supports (chips, beads, or glass slides), thus removing the barrier to comprehensive and repeated screening of sequences longer than that presently exemplified.

We gratefully acknowledge the financial support of the National Institutes of Health (CA41986, CA78045), the Department of Defense/Army (DAMD17-02-1-0135), the Skaggs Institute for Chemical Biology, and the award of an ACS Division of Medicinal Chemistry Graduate Fellowship sponsored by Wyeth-Ayerst (W.C.T., 2002–2003). W.C.T. is a Skaggs Fellow.

References

(1) Browne, M. J.; Thurlbey, P. L. *Genomes, Molecular Biology and Drug Discovery*; Academic: London, 1996.
(2) Matteucci, M. D.; Wagner, R. W. In pursuit of antisense. *Nature* **1996**, *384*, 20–22.
(3) Neidle, S.; Thurston, D. E. In *New Targets for Cancer Chemotherapy*; Kerr, D. J., Workman, P., Eds.; CRC: Boca Raton, FL, 1994.
(4) Thurston, D. E. Nucleic acid targeting: therapeutic strategies for the 21st century. *Br. J. Cancer* **1999**, *80* (Suppl. 1), 65–85.
(5) Choo, Y.; Sanchez-Garcia, I.; Klug, A. In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature* **1994**, *372*, 642–645.
(6) Neidle, S. Recent developments in triple-helix regulation of gene expression. *Anticancer Drug Des.* **1997**, *12*, 433–442.
(7) Trauger, J. W.; Baird, E. E.; Dervan, P. B. Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **1996**, *382*, 559–561.
(8) Knudsen, H.; Nielsen, P. E. Antisense properties of duplex- and triplex-forming PNAs. *Nucleic Acids Res.* **1996**, *24*, 494–500.
(9) Werstuck, G.; Green, M. R. Controlling gene expression in living cells through small molecule–RNA interactions. *Science* **1998**, *282*, 296–298.
(10) Chiang, S. Y.; Azizkhan, J. C.; Beerman, T. A. A comparison of DNA-binding drugs as inhibitors of E2F1- and Sp1-DNA complexes and associated gene expression. *Biochemistry* **1998**, *37*, 3109–3115.
(11) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. Regulation of gene expression by small molecules. *Nature* **1997**, *387*, 202–205.
(12) Dervan, P. B. Design of sequence-specific DNA-binding molecules. *Science* **1986**, *232*, 464–471.
(13) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. DNA affinity cleaving. Sequence specific cleavage of DNA by distamycin–EDTA–iron(II) and EDTA–distamycin–iron(II). *Tetrahedron* **1984**, *40*, 457–465.
(14) Galas, D. J.; Schmitz, A. DNase footprinting: a simple method for the detection of protein–DNA binding specificity. *Nucleic Acids Res.* **1978**, *5*, 3157–3170.
(15) Royer-Pokora, B.; Gordon, L. K.; Haseltine, W. A. Use of exonuclease III to determine the site of stable lesions in defined sequences of DNA: the cyclobutane pyrimidine dimer and cis and trans dichlorodiammine platinum II examples. *Nucleic Acids Res.* **1981**, *9*, 4595–4609.
(16) Van Dyke, M. W.; Hertzberg, R. P.; Dervan, P. B. Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA-Fe(II). *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5470–5474.
(17) Kuwabara, M. D.; Sigman, D. S. Footprinting DNA–protein complexes in situ following gel retardation assays using 1,10-phenanthroline-copper ion: *Escherichia coli* RNA polymerase-lac promoter complexes. *Biochemistry* **1987**, *26*, 7234–7238.
(18) Tullius, T. D.; Dombroski, B. A.; Churchill, M. E.; Kam, L. Hydroxyl radical footprinting: a high-resolution method for mapping protein–DNA contacts. *Methods Enzymol.* **1987**, *155*, 537–558.
(19) Hayashibara, K. C.; Verdine, G. L. Template-directed interference footprinting of cytosine contacts in a protein–DNA complex: potent interference by 5-aza-2′-deoxycytidine. *Biochemistry* **1992**, *31*, 11265–11273.
(20) Boger, D. L.; Johnson, D. S. CC-1065 and the duocarmycins: Understanding their biological function through mechanistic studies. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1438–1474.
(21) Rajski, S. R.; Williams, R. M. DNA Cross-linking agents as antitumor drugs. *Chem. Rev.* **1998**, *98*, 2723–2796.
(22) Tao, Z. F.; Saito, I.; Sugiyama, H. Highly cooperative DNA dialkylation by the homodimer of imidazole-pyrrole diamide–CPI conjugate with vinyl linker. *J. Am. Chem. Soc.* **2000**, *122*, 1602–1608.

(23) Hardenbol, P.; Wang, J. C.; Van Dyke, M. W. Identification of preferred distamycin–DNA binding sites by the combinatorial method REPSA. *Bioconjugate Chem.* **1997**, *8*, 617–620.
(24) Chaltin, P.; Borgions, F.; Van Aerschot, A.; Herdewijn, P. Comparison of library screening techniques used in the development of dsDNA ligands. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 47–50.
(25) Guelev, V. M.; Harting, M. T.; Lokey, R. S.; Iverson, B. L. Altered sequence specificity identified from a library of DNA-binding small molecules. *Chem. Biol.* **2000**, *7*, 1–8.
(26) Hamy, F.; Albrecht, G.; Florsheimer, A.; Bailly, C. An ARE-selective DNA minor groove binder from a combinatorial approach. *Biochem. Biophys. Res. Commun.* **2000**, *270*, 393–399.
(27) Larson, C. J.; Verdine, G. L. In *Bioorganic Chemistry: Nucleic Acids*; Hecht, S. M., Ed.; Oxford University: Oxford, 1996; pp 324–346.
(28) Blackwell, T. K.; Kretzner, L.; Blackwood, E. M.; Eisenman, R. N.; Weintraub, H. Sequence-specific DNA binding by the c-Myc protein. *Science* **1990**, *250*, 1149–1151.
(29) Blackwell, T. K.; Weintraub, H. Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* **1990**, *250*, 1104–1110.
(30) Thiesen, H. J.; Bach, C. Target Detection Assay (TDA): a versatile procedure to determine DNA binding sites as demonstrated on SP1 protein. *Nucleic Acids Res.* **1990**, *18*, 3203–3209.
(31) Chodosh, L. A.; Carthew, R. W.; Sharp, P. A. A single polypeptide possesses the binding and transcription activities of the adenovirus major late transcription factor. *Mol. Cell Biol.* **1986**, *6*, 4723–4733.
(32) Boger, D. L.; Fink, B. E.; Brunette, S. R.; Tse, W. C.; Hedrick, M. P. A simple, high-resolution method for establishing DNA binding affinity and sequence selectivity. *J. Am. Chem. Soc.* **2001**, *123*, 5878–5891.
(33) Boger, D. L.; Fink, B. E.; Hedrick, M. P. Total synthesis of distamycin A and 2640 analogs: A solution-phase combinatorial approach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative DNA binding affinity or DNA binding sequence selectivity. *J. Am. Chem. Soc.* **2000**, *122*, 6382–6394.
(34) Lee, L. G.; Chen, C. H.; Chiu, L. A. Thiazole orange: a new dye for reticulocyte analysis. *Cytometry* **1986**, *7*, 508–517.
(35) Nygren, J.; Svanvik, N.; Kubista, M. The interactions between the fluorescent dye thiazole orange and DNA. *Biopolymers* **1998**, *46*, 39–51.
(36) LePecq, J. B.; Paoletti, C. A fluorescent complex between ethidium bromide and nucleic acids. Physical–chemical characterization. *J. Mol. Biol.* **1967**, *27*, 87–106. Reviews: Morgan, A. R.; Lee, J. S.; Pulleyblank, D. E.; Murray, N. L.; Evans, D. H. Ethidium fluorescence assays. Part I. Physiochemical studies. *Nucleic Acids Res.* **1979**, *7*, 547–569. Morgan, A. R.; Evans, D. H.; Lee, J. S.; Pulleyblank, D. E. Ethidium fluorescence assay. Part II. Ezymatic studies and DNA–protein interactions. *Nucleic Acids Res.* **1979**, *7*, 571–594.
(37) For a complete list of references, see ref 32, footnote 9.
(38) Ethidium possesses relatively little sequence preference, although it displays a slight bias toward binding GC-rich DNA tracts. Each hairpin is normalized to its own 100% fluorescence to ensure that any bias in ethidium binding does not affect assay results, and accurate results may be obtained with any sequence composition.
(39) Boger, D. L.; Tse, W. C. Thiazole orange as the fluorescent intercalator in a high-resolution FID assay for determining DNA binding affinity and sequence selectivity of small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2511–2518.
(40) Broude, N. E. Stem–loop oligonucleotides: a robust tool for molecular biology and biotechnology. *Trends Biotechnol.* **2002**, *20*, 249–256.
(41) For the 5-bp variable region hairpin, the UV A_{260} at 25 °C may be converted to an accurate concentration by adjusting the millimolar extinction coefficients published by Invitrogen Corp. for ssDNA. Correction factors relating A_{260} at 90 °C/25 °C were determined to be 1.18, 1.13, 1.12, 1.12, 1.10, and 1.06 for the six combinations of 5-bp A/T sites contained in the 512 library: 5-bp AT, 4-bp AT/1 bp G/C, 3-bp AT/2-bp GC, 2-bp AT/3-bp GC, 1-bp AT/4-bp GC, and 5-bp GC, respectively.
(42) Browne, K. A.; He, G. X.; Bruice, T. C. Microgonotropens and their interactions with DNA. 2. Quantitative evaluation of equilibrium constants for 1:1 and 2:1 binding of dien-microgonotropen-a, -b, and -c as well as distamycin and Hoechst-33258 to d(GGCGCAAATTTGGCGG)/d(CCGCCAAATTTGCGCC). *J. Am. Chem. Soc.* **1993**, *115*, 7072–7079.
(43) Satz, A. L.; Bruice, T. C. Synthesis of fluorescent microgonotropens (FMGTs) and their interactions with dsDNA. *Bioorg. Med. Chem.* **2000**, *8*, 1871–1880.
(44) Scatchard, G. *Ann. N.Y. Acad. Sci.* **1949**, *51*, 660.
(45) Tse, W. C.; Ishii, T.; Boger, D. L. Comprehensive high-resolution analysis of hairpin polyamides utilizing a fluorescent intercalator displacement (FID) assay. *Bioorg. Med. Chem.* **2003**, *11*, 4479–4486.
(46) Hawkins, C. A.; Pelaez de Clairac, R.; Dominey, R. N.; Baird, E. E.; White, S.; Dervan, P. B.; Wemmer, D. E. Controlling binding orientation in hairpin polyamide DNA complexes. *J. Am. Chem. Soc.* **2000**, *122*, 5235–5243.
(47) Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. Influence of a terminal formamido group on the sequence recognition of DNA by polyamides. *J. Am. Chem. Soc.* **2002**, *124*, 2153-2163.
(48) Parks, M. E.; Baird, E. E.; Dervan, P. B. Optimization of the hairpin polyamide design for recognition of the minor groove of DNA. *J. Am. Chem. Soc.* **1996**, *118*, 6147–6152.
(49) Boger, D. L.; Dechantsreiter, M. A.; Ishii, T.; Fink, B. E.; Hedrick, M. P. Assessment of solution-phase positional scanning libraries based on distamycin A for the discovery of new DNA binding agents. *Bioorg. Med. Chem.* **2000**, *8*, 2049–2057.
(50) Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
(51) Woods, C. R.; Ishii, T.; Wu, B.; Bair, K. W.; Boger, D. L. Hairpin versus extended DNA binding of a substituted β-alanine linked polyamide. *J. Am. Chem. Soc.* **2002**, *124*, 2148–2152.
(52) Woods, C. R.; Ishii, T.; Boger, D. L. Synthesis and DNA binding properties of iminodiacetic acid-linked polyamides: characterization of cooperative extended 2:1 side-by-side parallel binding. *J. Am. Chem. Soc.* **2002**, *124*, 10676–10682.
(53) Wade, W. S.; Mrksich, M.; Dervan, P. B. Design of peptides that bind in the minor groove of DNA at 5′-(A,T)G(A,T)C(A,T)-3′ sequences by a dimeric side-by-side motif. *J. Am. Chem. Soc.* **1992**, *114*, 8783–8794.
(54) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7586–7590.
(55) Wade, W. S.; Mrksich, M.; Dervan, P. B. Binding affinities of synthetic peptides, pyridine-2-carboxamidonetropsin and 1-methylimidazole-2-carboxamidonetropsin, that form 2:1 complexes in the minor groove of double-helical DNA. *Biochemistry* **1993**, *32*, 11385–11389.
(56) Travis, A.; Amsterdam, A.; Belanger, C.; Grosschedl, R. LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function. *Genes Dev.* **1991**, *5*, 880–894.
(57) Waterman, M. L.; Jones, K. A. Purification of TCF-1 alpha, a T-cell-specific transcription factor that activates the T-cell receptor C alpha gene enhancer in a context-dependent manner. *New Biol.* **1990**, *2*, 621–636.
(58) Waterman, M. L.; Fischer, W. H.; Jones, K. A. A thymus-specific member of the HMG protein family regulates the human T cell receptor C alpha enhancer. *Genes Dev.* **1991**, *5*, 656–669.
(59) van de Wetering, M.; Oosterwegel, M.; Dooijes, D.; Clevers, H. Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *Embo J.* **1991**, *10*, 123–132.
(60) Grosschedl, R.; Giese, K.; Pagel, J. HMG domain proteins—Architectural elements in the assembly of nucleoprotein structures. *Trends Genet.* **1994**, *10*, 94–100.
(61) Eastman, Q.; Grosschedl, R. Regulation of LEF-1/TCF transcription factors by Wnt and other signals. *Curr. Opin. Cell. Biol.* **1999**, *11*, 233–240.
(62) Aoki, M.; Hecht, A.; Kruse, U.; Kemler, R.; Vogt, P. K. Nuclear endpoint of Wnt signaling: Neoplastic transformation induced by transactivating lymphoid-enhancing factor 1. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 139–144.
(63) Giese, K.; Amsterdam, A.; Grosschedl, R. DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. *Genes Dev.* **1991**, *5*, 2567–2578.
(64) Love, J. J.; Li, X. A.; Case, D. A.; Giese, K.; Grosschedl, R.; Wright, P. E. Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **1995**, *376*, 791–795.
(65) Ham, Y. W.; Tse, W. C.; Boger, D. L. High-resolution assessment of protein DNA binding affinity and selectivity utilizing a fluorescent intercalator displacement (FID) assay. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3805–3807.
(66) Yeung, B. K. S.; Tse, W. C.; Boger, D. L. Determination of binding affinities of triplex forming oligonucleotides using a fluorescent intercalator displacement (FID) assay, *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3801–3804.

AR030113Y